Towards the improvement of food flavor analysis through the modelling of olfactometry data and expert knowledge integration

Alice Roche

PHD THESIS

For the degree of *Docteur de l'Université de Bourgogne Franche-Comté*

Ecole Doctorale Environnement-Santé

Discipline: Food science

Presented by

Alice Roche

TOWARDS THE IMPROVEMENT OF FOOD FLAVOR ANALYSIS

THROUGH THE MODELLING OF OLFACTOMETRY DATA AND EXPERT KNOWLEDGE INTEGRATION

Defended the 25th of October 2018 in Dijon, France

Composition of the Jury:

Pr. Catherine Dacremont	AgroSup Dijon	Chairman
Dr. Didier Dubois	CNRS Toulouse	Reviewer
Pr. Vicente Ferreira	University of Zaragoza	Reviewer
Dr. Nicolas Monmarché	Université de Tours	Examiner
Dr. Sophie Tempère	Université de Bordeaux - ISVV	Examiner
Dr. Joel Mainland	Monell Chemical Senses Center	Invited
Dr. Nathalie Perrot	INRA - GMPA	PhD supervisor
Dr. Thierry Thomas-Danguin	INRA - CSGA	PhD supervisor

Roses are red.

Violets are blue.

Let's pour a glass of wine,

and read this thesis through.

ACKNOWLEDGMENTS

This manuscript contains the results of three years of research in Food science and more precisely on Flavors. Three years of PhD those were motivated by former experiences. Indeed, we have to shortly go back in time to understand why I am writing my thesis today. I began to study flavor perception four years ago at the Food science laboratory of Cornell University in Geneva (NY, USA) with Pr. Terry Acree. I would like to first thank him in addition to Ed Lanvin and Steve Wickoff. They taught me a lot about flavor science, flavor analysis and olfactometers. Motivated to continue my work on Flavors, one year later I did an internship as a Food Technologist at Givaudan, in the "Sweet aromas Research & Development" team, where I discovered the industrial side of the field. With the desire to learn more and to conduct a long-term project about food flavor, I did this PhD, my PhD. I am proud to have written this manuscript and I would like to thank here all the people who supported me during this journey and who are contributing to who I am today.

For their expertise, I thank Dr. Didier Dubois and Pr. Vicente Ferreira for their role as external reviewers. I also thank Pr. Catherine Dacremont, Dr. Nicolas Monmarché and Dr. Sophie Tempère for their services during the defense examination process. To my committee members, Dr. Elisabeth Guichard, Dr. Evelyne Lutton, and Pr. Evelyne Vigneau, I warmly thank you for your insights, advices, and support throughout the development of my PhD project.

I would also like to thank the former and actual director of the Center for Taste and Feeding Behavior (CSGA), respectively Luc Pénicaud and Lionel Bretillon, as well as Elisabeth Guichard and Christian Salles to have welcomed me in their team.

I am thankful to the funding I receive during my PhD, the French National Institute for Agricultural Research (INRA) and the Conseil Régional de Bourgogne for my research allowance to enable me to conduct my research in France and abroad, the Ecole Doctorale Environnement Santé of the Burgundy University for its support to travel abroad to conduct research and present my results in conventions and the European Chemoreception Research Organization (ECRO) for the travel grant ECRO 2018 award which enabled me to attend the ECRO annual meeting in Würzburg (Germany).

For their guidance and support throughout this work, I express my thanks and deepest gratitude to my supervisors Thierry Thomas-Danguin and Nathalie Perrot. I will always be grateful for what I have learned from you. We share a lot of science but also good moments together.

Now let's go to the other side of the Atlantic Ocean to thank Joel Mainland for the warm welcome in his lab at the Monell Chemical Senses Center in Philadelphia. Thank you for your precious advices.

Back to the French side, a big special thanks to Noëlle, Baptiste and again Thierry with whom I shared the office. We shared more than that: creative and musical moments, laughs, hs-hf recipes… Thierry, aka Master ThiThoDan, it was a pleasure to share running moments as well as Game of Thrones recaps… Noëlle, I will miss you a lot, I liked so much the way we supported each other in our work and our running goals. Thank you for your kindness and support. Baptiste, Baptou, El Bapto… what to say exactly but #INRAIZIIII? Long story short, I can resume our time together by Work hard - Play hard, or the other way around in fact. I would like to thank Sébastien, my predecessor, to have been there at the very beginning and fed the office with muffins. Thanks to Charlotte for her help as well as the ideal guided tour of Dresden. Thomas (Thomas Alwaysmore), you have your place here also for the good times and your unrivaled sense of humor. Thank you also to all the trainees I met during these three years (kisses to you Kamilia).

Even if I did not spend that much time in Grignon, I do not forget my PhD-mate Thomas and Marc. Thank you for your advices on the modelling parts. Good luck to both of you.

I would like to thank the Mainland lab and especially Casey, Nicolle, and Marissa who took care of me in Philadelphia: concerts, pop-up beer gardens, 4th July, Halloween… You contributed more than you think to my enjoyment of my stay. John, I still miss you a lot. You were the best coconut-head-office-mate and I am so glad to have spent six months with you (are you still trying to learn some French? If not: Booo). At the Monell I also met two amazing girls: Federica and Barbara (the Euro team). Fede, thank you for everything! With you I learned how to pronounce Baiocchi, we translated Despacito song, we ran a half-marathon and finished by our victory walk… *Grazie mille*.

During my PhD I was selected as other PhD students to be part of the Agreenium international research facility program, EIR-A, and I would like to thank here the extraordinary people I met: the Auriculus team. Since our first time in Montpellier, we are supporting each other through the big adventure which a PhD is.

Among this extraordinary people I met Nathan. Thank you so much for everything: listening to me when I taught you about flavors, your support, walking together along the seaside... You were more than perfect. I still cannot realize that we met during our PhDs; you are the hidden gift of this experience. Ps: Keep going, you are next!

To my closest friends, thank you for your patience, because even if we saw each other less than usual, nothing changed. To Christelle, my Antidable, my lightenin' strike, I know we already have a lot of memories together but there will be more. A big hug to Ghislain, also called Gisou since more than 16 years. Clément, Florent, Lucile, Vivien, thank you for all the moments we shared: Ultimate, kebab, beers, concerts, games nights… The life in Dijon was so good with you because even when it was cold, you were there to warm the atmosphere.

To my family, *merci*! Nothing would have been possible without you. You make me feel strong and positive at all times and you are always there wherever I am (Dijon, Paris, Geneva, Philadelphia). Pap and Mam, thank you for everything: your unconditional support, the breaks in Valence with delicious food and the famous Sunday's walk… Nicolas, aka Marsouin or Marsupial according to the days, I am so proud of you! Thank you for being the best brother I could have imagined. *Papi Romain et Mamie Gisèle, merci pour tous les moments que nous avons passé ensemble. Vous avoir eu proches de moi pendant mes 8 ans d'études à Dijon a contribué à ma réussite.*

Thank again to all of you.

Merci. Grazie. Gracias. Danke.

CONTENTS

FOREWORD

Food choice, acceptance, and consumption are not determined solely by physiological and nutritional needs. Some of the other factors that influence food choice in a conscious or unconscious manner include: economical, psychological, environmental, and cultural determinants. Flavor is also consistently reported as having a major influence on food behavior along with pleasantness. Flavor is a sensory percept induced by food or beverage tasting and relies on the integration of taste, smell and somatosensory stimuli. Working on food flavor perception should contribute to the creation of food products in agreement with consumers' expectations and nutritional recommendations. To do so, a key goal is to gain insight into the link between the composition of food and its flavor and to develop relevant models with efficient predictive abilities.

Among the sensory modalities involved in food flavor perception, this manuscript focuses on odor perception (smell). Food odor analysis usually relies on the quantification and identification of a list of odorants. Several approaches attempted to predict the odor of odorants on the basis of their molecular structure. Such methods focused on single odorants; however odors we perceive in every-day life are mixtures including many different odorants at varying concentrations. Therefore, the odor analysis process is not sufficient to predict the odor profile of a given food product since the perception of odorants' mixture is not the simple sum of the odor of each odorants but the results of numerous and complex perceptual interactions.

In order to improve the food odor analysis efficiency, there is a need to take into account the perception of odorants' mixture to predict the odor profile of food products. To do so, the presented work explored two strategies. The first one is based on structure-odor relationships and aims to predict the similarity among odorants' mixture on the basis of their molecular structure. The second one relies on an innovative strategy which combines food odor analysis with expert knowledge on aroma formulation through a modelling approach relying on fuzzy logic and optimization. This multidisciplinary thesis project combines methodologies from the modelling to the food science, chemical, and sensory disciplines.

The developed strategies have been applied to the odorant composition of red wines, as a good example of a complex food matrix. Indeed, wine is a widely studied product in terms of odor/aroma. Moreover, the project was conducted in Dijon, capital of Burgundy known for its *Grand cru* wines: "the vines of France and milk of Burgundy" (King Lear, Shakespeare).

The thesis manuscript contains five chapters.

The review of the literature (Chapter 1) firstly gives a definition of flavor. Then, odorants' mixture perception is highlighted through examples of interactions between odorants at the perceptual level. Food odor measurements are then presented and are followed by the review of the predictive approaches applied to odor perception. Next, because expertise is integrated into the second strategy presented, the focus is set on food odor expertise. The chapter is concluded by a critical conclusion of the main literature items, followed by the presentation of the thesis core research questions and the scientific strategy of the thesis work.

Chapter 2 details the methodologies used in this work. First, the wines used in the predictive approaches are characterized through analytical and sensory analysis. Secondly, the collection of expertise from flavorists is described. Then, the concept of ontologies is explained. Lastly, fuzzy logic is defined and an example of application is detailed.

In the three following chapters (Chapter 3, Chapter 4, and Chapter 5), the work achieved during the thesis is presented through three articles.

In Chapter 3, the first predictive modelling strategy is presented. The link between odorants' molecular structure and mixture similarity is established for complex mixtures, including real food matrices (e.g. wines).

In Chapter 4, the construction of an ontology for odor perceptual space is presented. This knowledge representation relies on semantic odor information in part collected from expert flavorists. The ontology is used in a predictive approach to predict odor quality perception from complex mixtures, namely wines. It will be implemented in the following chapter as a basis of an expert system.

Chapter 5 presents the second predictive strategy which combines ontology, fuzzy logic and optimization methodologies to form an operational expert system. The developed model is able to predict the odor profile of wines qualitatively and quantitatively through the estimation of the intensity of odor sensory attributes.

The manuscript ends with a general discussion about the work achieved, the perspectives addressed for further research and a general conclusion.

CHAPTER 1: LITERATURE REVIEW

STATE OF THE ART

I. Flavor

Flavor is a multi-modal perception arising from the consumption of food and beverage (Thomas-Danguin, 2009). This perception results from the integration of simultaneous information transmitted by different sensory systems: gustatory, somatosensory and olfactory (Laing and Jinks, 1996). However, there continues to be a significant debate over just how many modalities should be included in flavor construction (Small, 2012): chemical sensations that arise from the mouth (gustatory and olfactory receptors), smell (olfactory receptors), sound (sound perception), texture (tactile receptors), and sight (visual perception)? In this manuscript, the flavor perception is defined as the one created from the activation of the chemical senses: taste, somatosensation (nasal and oral) and smell (Figure 1.1).

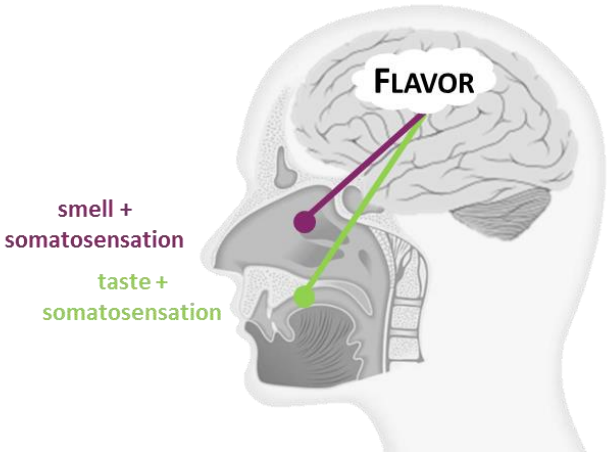


Figure 1.1: Modalities of flavor perception. Flavor is the result of the integration of simultaneous information transmitted by the gustatory, somatosensory and olfactory systems.

1. Gustation

Gustation refers to the sense of taste, corresponding to the sensation produced when tastants interact with taste receptors of taste receptor cells located on taste buds in the oral cavity, mainly on the tongue but also on the soft palate, and in the pharyngeal and the laryngeal regions of the throat (Breslin & Huang, 2006). Taste compounds belong to different chemical classes such as mineral salts, organic acids, amino acids, nucleotides, carbohydrates, peptides, proteins, and some phenolic or heteroatomic compounds (Thomas-Danguin et al., 2012; Briand & Salles, 2016).

Our sense of taste allows us to detect five basic qualities: bitter, salty, sour, sweet, and umami. A sixth quality has recently been proposed regarding the ability to taste fatty acids (Mattes, 2010). Astringency (Jiang et al., 2014), kokumi (mouthfullness) (Maruyama et al., 2012), and starchiness (Lapis et al., 2016) are also studied and might be added to the taste range.

This sensory modality allows the evaluation of food and beverage in order to accept it by swallowing or reject it by expectoration. In addition, taste has a clear role in nutrient sensing and is thus highly involved in eating behavior (Boesveldt & de Graaf, 2017).

2. Somatosensation

Somatosensation refers to the process that conveys information regarding the body surface and its interaction with the environment. It can be subdivided into mechanoreception, thermosensation, and nociception. During food consumption, somatosensation is caused by the excitation of the trigeminal nerve by trigeminal compounds in the mouth or in the nose. The resulting perceptions are: spiciness (or pungency), astringency, burning or cooling. Trigeminal compounds can include taste or odor compounds as for example menthol (CAS 1490-04-6) which induces a cooling sensation as well as a minty odor percept (Salles et al., 2012).

3. Olfaction

Olfaction refers to the sense of smell which also largely impacts eating behavior. In most cases, food odor reflects food identity and typicality. For instance, it has been demonstrated that odor exposure induces appetite specifically for the cued food (Boesveldt & de Graaf, 2017).

a. Odorants and odors

There is a general agreement that molecules giving rise to an odor sensation should be referred to as odorants, however the usage of the term odor is less clear (Hudson, 2000). Thus, odorants are molecules capable of being translated by the olfactory nervous systems into odor percepts. As mentioned earlier, odorants can also stimulate the activation of the nasal trigeminal system (Doty et al., 1978).

Odorants, as molecules, are objectively definable in terms of their physico-chemical characteristics. Odorants are typically organic volatile compounds of low molecular weight (<400 Dalton), they may be aliphatic or aromatic, may be saturated or unsaturated, and may have any of several polar functional groups. Molecular structures of odorants are available on different databases, such as PubChem (http://pubchem.ncbi.nlm.nih.gov/). In addition, with the advent of software such as Dragon (Talete, Milan, Italy), it is now possible to calculate a wide number of structural parameters (more than 4800) from the molecular structure of odorants.

Odorants are also defined with their psychophysical characteristics such as detection threshold. Odor detection threshold corresponds to the minimum concentration at which 50% of a human panel can detect the presence of an odor without being able to characterize or recognize the stimulus. Thresholds of odorants differ by many orders of magnitude. For example, the thresholds in water for bis(2-methyl-3-furyl)disulphide (CAS 28588-75-2) and ethanol (CAS 64-17-5) are respectively $0.32x10^{-9}$ and 0.99 $g.L^{-1}$ (Czerny et al., 2008). However it has to be noted that thresholds in literature vary widely due to the possible contamination with other odorants, and threshold determination methods (Chastrette, 1998).

Furthermore, odorants are characterized by their intensity as a function of concentration. This relationship is not a simple binary function but a psychometric unique function for each odorant (Delahunty et al., 2006). Psychometric functions can best be considered as sigmoidal shape in a plot of log concentration against perceived intensity. Modelling attempts to predict

the intensity of odorants from a wide range of concentration concluded that models derived from the Hill equation should be preferred (Chastrette et al., 1998).

Odorants can also be described with the odor percept they evoke, but we will see further that naming odors can be a complicate task and often lead to a so-called "tip-of-the-nose" phenomenon (Lawless & Engen, 1977). Recently, it has been estimated that humans can detect and discriminate more than one trillion different odorants (Bushdid et al., 2014), though the exact number remains unknown. Thus we can assume an even wider number of odor percepts.

b. Orthonasal and retronasal olfaction

Orthonasal olfaction refers to the route of sensing odors in our environment involving inhalation of odorants through the nares. However, odorants can also enter the nose from the mouth through internal nares in the upper palate during food consumption via the retronasal route (Delime et al., 2016) (Figure 1.2). This second mechanism of odorants perception only occurs when the velum tongue border is open to allow odorants released from food into the mouth to be transported by the swallowing breath into the nasal cavity (Buettner et al., 2001).

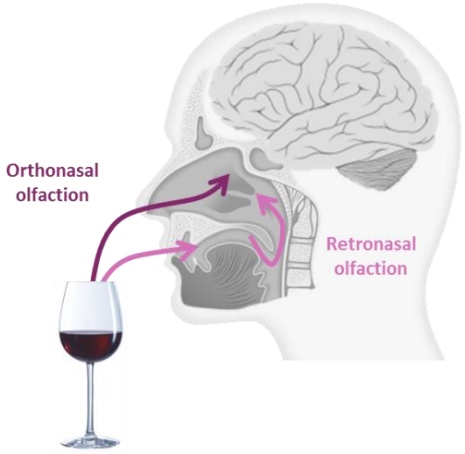


Figure 1.2: Schematization of ortho- and retronasal olfaction pathways.

Different terminologies are commonly used depending on the way molecules are involved in olfaction. Odor percept is used when odorants are stimulating the olfactory system via the orthonasal pathway whereas aroma percept is used when odorants are stimulating the

olfactory system via the retronasal pathway. Thus odor compounds and aroma compounds are the same compounds; the difference relies on the olfactory route and their processing by the brain (Heilmann & Hummel, 2004). We will further use only the term odorant for odor-active molecules; however the aroma terminology might be used according to the literature context.

Perception via the ortho- and retronasal pathways have been shown to be different for the same odorants and odorants' mixtures (Goldberg et al., 2018). Indeed, there are differences in the perception of the duration of odors. For example, the perceived intensity of the lemon odor of citral and the vanilla odor of vanillin persisted longer by retronasal olfaction (Kuo et al., 1993). Furthermore, it seems that different neural processes occur according to the olfactory pathway (Small et al., 2005; Ishii et al., 2008).

c. Mechanisms of olfaction

Now that we have defined the concepts of odors, odorants and the differences between orthonasal and retronasal olfaction, the next step is to understand how odorants elicit odor percepts. Although we focus on the human olfactory system, studies which have provided information about the olfactory system structure and cellular functioning were done mostly on animals and especially rodents.

Odorants enter the nasal cavity either via the nose, by sniffing (orthonasal olfaction), or via the mouth, during eating or drinking (retronasal olfaction). Then odorants dissolve in the mucus that lines the superior portion of the cavity and are detected by the olfactory receptors (OR) located on the dendrites of olfactory sensory neurons (OSN).

Once an OR has been activated, a cascade of events is initiated transforming the chemical-structural information contained in the stimulus into a neural signal, i.e. a membrane potential. This signal is projected to a first relay in the brain, the olfactory bulb, from where it is transmitted to the olfactory cortex and higher regions of the brain. This activity pattern is then decoded as a particular odor percept.

Each OSN expresses a single olfactory receptor gene (Monahan & Lomvardas, 2015; Nagai et al., 2016). As shown in Figure 1.3, axons from OSN expressing the same OR gene converge in the same regions of the olfactory bulb, so that each OR gene has a corresponding glomerulus in the olfactory bulb (Ressler et al., 1994; Vassar et al., 1994; Mombaerts et al., 1996).

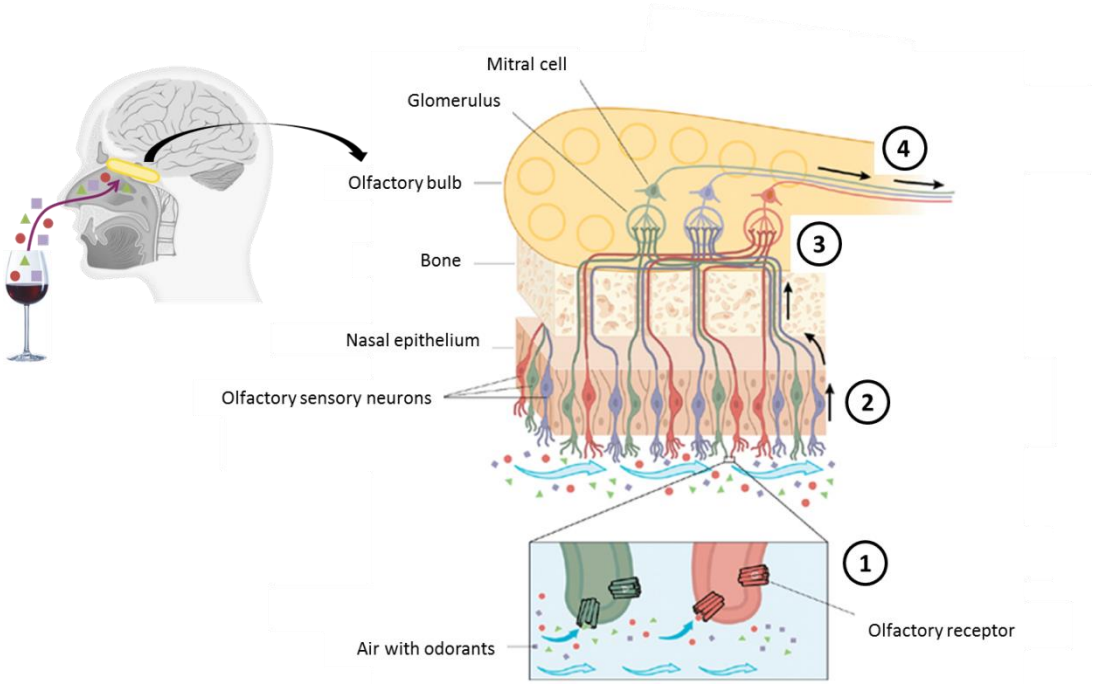


Figure 1.3: Mechanisms of olfaction. 1. Odorants enter in the nasal cavity and bind to olfactory receptors (OR) on the dendrites of olfactory sensory neurons (OSN). 2. Olfactory receptor cells, i.e. OSN, are activated and transduce signals, which are sent to the olfactory bulb. 3. Once in glomeruli, signals are transmitted to second order neurons in the olfactory bulb. 4. Signals are then heading to the olfactory cortex and other regions of the brain to create an odor percept. Figure adapted from Rinaldi (2007).

A given odorant activate a specific group of glomeruli in the olfactory bulb and the resulting signal lead to odorant discrimination (Duchamp et al., 1974). The human genome contains ~400 OR functional genes (Zozulya et al., 2001; Malnic et al., 2004) which belong to the large gene family of G protein-coupled receptors (Buck & Axel, 1991). Each OR is thought to recognize more than one odorant and each odorant can be recognized by more than one OR so that each odorant activates a specific combination of OR (Figure 1.4) (Malnic et al., 1999).

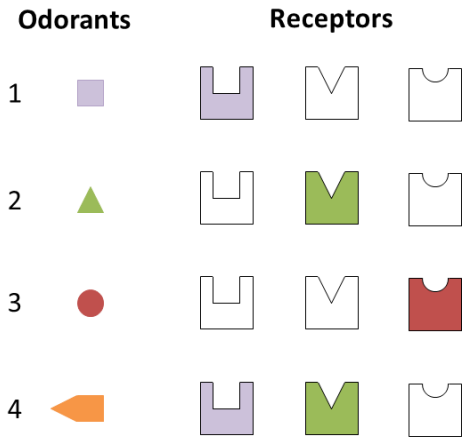


Figure 1.4: Combinatorial coding of odorants. Olfactory receptors (OR) colored are activated by the odorant on the left. Each odorant has its own receptor's code. The wide number of possible combinations explains the ability of the olfactory system to discriminate a wide number of odorants. Figure adapted from Malnic et al. (1999).

Olfactory coding mechanisms depend on which odorant activates a given OR, and at which concentration. On the one hand, the existence of many OR and the wide number of possible ligands makes the deorphanization of these receptors a challenging task (Peterlin et al., 2014). So far only a very small fraction of the human OR has been linked to their ligands (Mainland et al., 2015). However, new techniques should contribute to the identification of human OR that recognize odorants of interest (Armelin-Correa & Malnic, 2017). On the other hand, changes in odorant concentration result in changes in the combination of receptors that recognize the odorant. Indeed, the increase of an odorant's concentration leads to the activation of additional OR (Malnic et al., 1999).

Thus, transduction of olfactory stimulation into odor percepts depends on various parameters such as the odorant's structure and its concentration-intensity curve. Moreover, we should not forget that sensitivity to single odorants varies greatly among subjects with a normal sense of smell (Keller et al., 2012) which is another issue in the prediction of OR activation and odor percepts elicited by odorants.

d. Verbalizing odors

Even though odorants are omnipresent in our daily lives, verbalizing odor percepts is not an easy task. If asked to name everyday odors like peanut butter, cinnamon or strawberry, most

people can only name half of them correctly (Cain, 1979). However, accurate description of odors is crucial for communication purposes in food, flavoring and fragrance industries.

Characterizing odors in verbal terms is a complex process that consists in identifying a set of descriptors that best synthesizes the olfactory perception. Generally, this semantic characterization results in a limited set of odor descriptors (Abe et al., 1989; Zarzo & Stanton, 2006; Iatropoulos et al., 2018), also called odor descriptions (Tromelin et al., 2017), odor notes (Chastrette et al., 1991), perceptual descriptors (Kumar et al., 2015), or flavor descriptors (Martínez-Mayorga et al., 2011). In this manuscript, we used the term odor descriptors to describe odors.

When asked to describe odors in verbal terms, neophytes provide sets of odor descriptors relying mainly on the source of the smell, as "odor of/the/a N" where N denotes an odor source (Rouby et al., 2005; Dubois, 2006), and focused on the stimulus intensity (weak or strong) and its pleasantness ("it smells good" or "it smells bad") (Vassiliadou & Lammert, 2011). Such terms are subjective and thus cannot be shared among individuals (Barkat-Defradas & Motte-Florac, 2016). It was also demonstrated that odor naming for neophytes is facilitated by forced choice among a list of descriptors. Such finding supports the hypothesis that odor naming relies more on global semantic discriminations than on lexical access (Rouby et al., 2005).

On the contrary, experts seem to put the hedonic tone of odor percepts aside when they characterize an olfactory perception (Sezille et al., 2014). Indeed, expert flavorists and perfumers learn a common language to calibrate the description/characterization of odors. We can mention the Champ des odeurs© (Jaubert et al., 1995) and Sense It™, a global flavor language at Givaudan (Veinand, 2015). Wheels of odors can also help trained sensory panelists or experts to share a common referential to qualify odors of specific food products (caramel: Paravisini et al., 2014; honey: International Honey Commission (IHC) http://www.ihc-platform.net/reports.html; wine: Noble et al., 1987).

Databases compiling odor descriptors for large sets of odorants are available (The good scents company: Luebke, 1980; Atlas of odor character profiles: Dravnieks, 1985; Arctander's handbook: Arctander, 1969; Fenaroli's handbook: Burdock, 2010; Flavornet: Arn & Acree, 1998; Flavor-Base: Leffingwell & Associates, http://www.leffingwell.com; Flavors and Fragrances of Sigma-Aldrich: http://www.sigmaaldrich.com/industries/flavors-and-fragrances.html). The number of odorants in each database and the methodology to obtain the

odor descriptors is detailed in Table 1.1. As an example, the odor description of the odorant Ethyl propionate (CAS 105-37-3) is presented according to the seven databases. This highlights the lack of consensus among the databases and of agreement about the number of descriptors essential to cover the complete range of odor stimuli which varies from 4 to 146 (Chastrette, 2002). Thus, the evaluation of what we could expect to be the most simple odor percept, which is related to a single odorant, is already highly variable.

It also important to mention that the odor percept elicited by an odorant may change according to its concentration (Gross-Isserof & Lancet, 1988). Indole, for example, has an "extremely diffusive and powerful odor, almost tarry-repulsive and choking" when concentrated but is perceived as floral and pleasant when diluted (Arctander, 1969). Such phenomenon may rely on peripheral processes as well central processing, including quality encoding and mnemonic devices (Cain, 1979).

Another issue not discussed here about odor verbalization is linked to cultural differences which can affect the ability to identify or describe odors (Chrea et al., 2004; Ferdenzi et al., 2016).

The flavor of food, along with its appearance and texture, is considered to be decisive for the consumer in the selection and ingestion of food (Fisher & Scott, 1997). Because flavor perception arises from the central integration of multiple sensory inputs (Small & Prescott, 2005), modalities contributing to flavor are commonly confused (Murphy & Cain, 1980; Rozin 1982; Chaudhari & Roper, 2010; Fondberg et al., 2018). Distinguishing the different modalities is still possible, especially when attention is drawn to particular sensory characteristics. Among the sensory dimensions involved in food flavor perception, the olfactory component is critical because it determines most of the time the identity and the typicality of the food, which drive its overall quality and recognition by consumers (Hornung & Enns, 1986).

Table 1.1: Comparison of seven odor descriptors databases. *: descriptors are presented along with the percent of subjects who used the given descriptor (only percentages of more than 20% are shown).

Databases	Number of odorants	Data collection	Odor description of Ethyl propionate (CAS 105-37-3)
The good scents company (Luebke, 1980)	2933	Odor description from one to several sources	Sweet, fruity, rum, juicy, fruit, grape, pineapple
Atlas of odor character profiles (Dravnieks, 1985)	138	120-140 panelists rated the smell of the odorants for 146 descriptors on a numeric scale (0-5)	Fruity, other than citrus (37.41); Fragrant (42.45); Aromatic (33.81); Sweet (52.52); Light (23.02) *
Arctander's handbook (Arctander, 1969)	3102	Odor description provided by Arctander himself	Very volatile, ethereal, fruity-rum-like odor
Fenaroli's handbook (Burdock, 2010)	More than 2000	Data collected from published articles and books	Ethyl propionate has an odor reminiscent of rum and pineapple
Flavornet (Arn & Acree, 1998)	738	Data collected from articles published since 1984 using GC-O to detect odorants in natural products	Fruit
Flavor-Base (Leffingwell & Associates, http://www.leffingwell.com)	About 4324	Data collected from published work	Ethereal; Fruity; Rum
Flavors and Fragrances of Sigma-Aldrich (https://www.sigmaaldrich.com/industries/flavors-and-fragrances.html)	More than 1500	Unknown	Grape; fruity; ethereal; rum; sweet; wine-like

II. Mixtures of odorants

We outlined that transduction of olfactory stimuli into odor percepts depends on many characteristics of single odorants. However, smelling single odorants in every-day life rarely occurs, instead we perceive complex mixtures of tens, hundreds or even thousands of odorants. For example, chemical aroma analyses of wines usually reveal about 60 to 80 odorants regardless of the wines' color (Aznar et al., 2001; Lee & Noble, 2003; Zhao et al., 2017).

Within the context of odorants' mixtures, the overlapping response profiles of olfactory receptors (OR) and thus olfactory sensory neurons (OSN) introduce the possibility of interactions which occur at the peripheral level of the olfactory system or its first relay in the brain and may further influence the perceived intensity of mixtures and the odor percepts they evoke (Berglund et al., 1976; Goyert et al., 2007).

1. Peripheral interactions

Peripheral interactions can occur during the first step of the olfactory process, when odorants bind to OR and the resulting activity pattern is transmitted to the olfactory bulb (OB) in the brain via the OSN.

a. Olfactory receptor (OR) level

Odorants activate different but sometimes overlapping subset of OR, which defines an activity pattern specific to one odorant. When odorants are mixed, a code for the mixture is expected to be the sum of the coding patterns of its components. However, there can be competitive interactions when odorants bind to the same receptors. For example in Figure 1.4, odorant 4 can bind to the same OR as the odorants 1 and 2.

This mechanism could involve either two agonist odorants, i.e. both odorants bind to the same receptor and activate it, or one agonist and one antagonist. An antagonist blocks or dampens the biological response by binding to and blocking a receptor rather than activating it (Spehr et al., 2003; Oka et al., 2004; Sanz et al., 2005; Jacquier et al., 2006). For example, the odorant bourgeonal (CAS 18127-01-0) is a powerful agonist for the human olfactory receptor

hOR17-4, while the odorant undecanal (CAS 112-44-7) fails to activate this receptor. However, when the two odorants are mixed, the hOR17-4 response is suppressed, which indicates that undecanal inhibits the receptor activation by bourgeonal (Spehr et al., 2003). Moreover the interaction identified at the OR level between bourgeonal and undecanal exhibits a strong inhibitory effect on bourgeonal odor at the perceptual level in humans (Brodin et al., 2009). Odorants could act both as agonist for some OR and as antagonist for others (Oka et al., 2004). As illustrated, when antagonism between odorants occurs, a smaller number of OR are activated than the additive number of active OR by single odorants (Figure 1.5). This dual function as agonist or antagonist to OR provides complexity in the encoding mechanism of an odorants' mixture at the receptors level (Rospars et al., 2008).

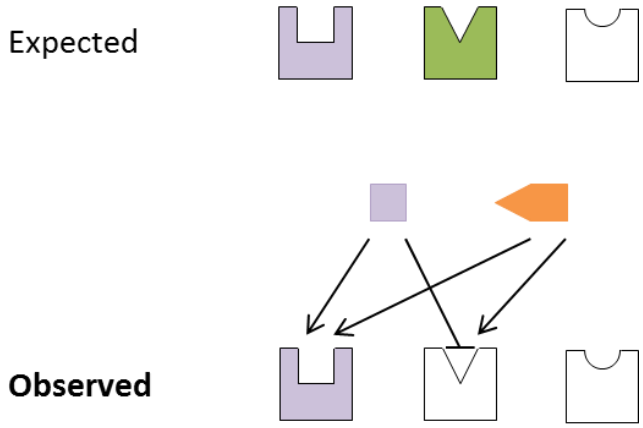


Figure 1.5: Combinatorial coding of odorants' mixture is not the sum of the code of each odorant. Olfactory receptors (OR) colored are activated. Taking the example of the odorants 1 and 4 in Figure 1.4, odorant 1 is an agonist for the violet receptor but an antagonist for the green receptor whereas odorant 4 is an agonist of the receptors violet and green.

b. Olfactory sensory neuron (OSN) responses

Different types of interactions were observed in rodents regarding the responses of OSN to binary mixtures and their components (Duchamp-Viret et al., 2003). The response of OSN to binary mixtures was observed to be lower, higher or equivalent to the highest single odorant response. Interactions also depend on the odorants included in the mixtures and their

concentration ratios because a shift between interaction types was reported as a function of odorant concentration (Duchamp-Viret et al., 2003; Chaput et al., 2012).

Interactions in rat OSN responses to binary mixtures of isoamyl acetate (CAS 123-92-2) and whiskey lactone (CAS 39212-23-2) were in concordance with the perceptual responses to the same mixtures in humans (Chaput et al., 2012). Indeed, rat OSN responses to the mixtures were enhanced or reduced depending on the concentration of whiskey lactone in the mixture. Similarly, in humans, the fruity intensity given by isoamyl acetate was increased by low concentrations of whiskey lactone and decreased by high concentrations.

c. Olfactory bulb (OB) responses

Each OR has a corresponding glomerulus in the olfactory bulb (Figure 1.3). Thus, several glomeruli are activated by OSN when OR are stimulated with odorants' mixtures. However, as for the OR, the number of glomeruli activated differs from the sum of the glomerulus activated by each odorant of the mixture (Bell et al., 1987; Lin et al., 2006; Grossman et al., 2008). This suggests that spatial activity patterns within the olfactory bulb are not sufficient to predict odorant recognition in mixtures.

Together these results suggest that mixtures' perception is reshaped at each level of processing and signal integration (OR, OSN, OB), odorants being encoded differently when they are stimulating the olfactory system alone or in mixtures. This also implies further processing which occur in higher brain areas (Boyle et al., 2009).

2. Perceptual interactions

Peripheral interactions shape the odor signal, which seems to determine the perceptual features of complex mixtures (Kay et al., 2003). The consequences of the non-linear integration of the chemical information carried by an odorants' mixture can indeed be observed at the level of odor perception, i.e. their quality, intensity and pleasantness.

a. Odor percept

The odor percept resulting from an odorants' mixture, as defined by Berglund et al. (1976), can be homogeneous when a single odor is perceived from the mixture or heterogeneous when several odors are perceived from the mixture.

A homogeneous percept is elicited when the odors of the mixed odorants lose their individual odor quality and blend into a new odor perceived as an entity, this is called odor blending. The perception of such mixtures is considered as configural or synthetic (Berglund & Olsson, 1993; Laing, 1994; Jinks & Laing, 2001; Kay et al., 2005). Overshadowing (Kay et al., 2005) and masking (Cain & Drexler, 1974) are also considered as homogeneous percepts because the odor quality of one component completely covers the odors of the other components.

A heterogeneous percept is elicited when at least some of the component odors can be perceived within the mixture. The perception of such mixtures includes analytical (Berglund & Olsson, 1993) or elemental (Kay et al., 2005) perception which occurs when the stimuli keep their individual qualities.

Figure 1.6 illustrates the perceptual interactions in binary mixtures. In the case of more complex mixtures, it has been suggested that the odor percept of the mixture is more frequently different from the odor qualities of their constituting odorants. In other words, complex mixtures are more inclined to evoke the perception of a new odor (Livermore & Laing, 1998; Ferreira, 2012b; Lindqvist et al., 2012). The case of partial blending highlights the fact that both mechanisms can occur at the same time. Indeed, in the case of complex mixtures, subjects are only able to identify three or four components, the remaining being perceived as a single entity or masked (Laing & Francis, 1989).

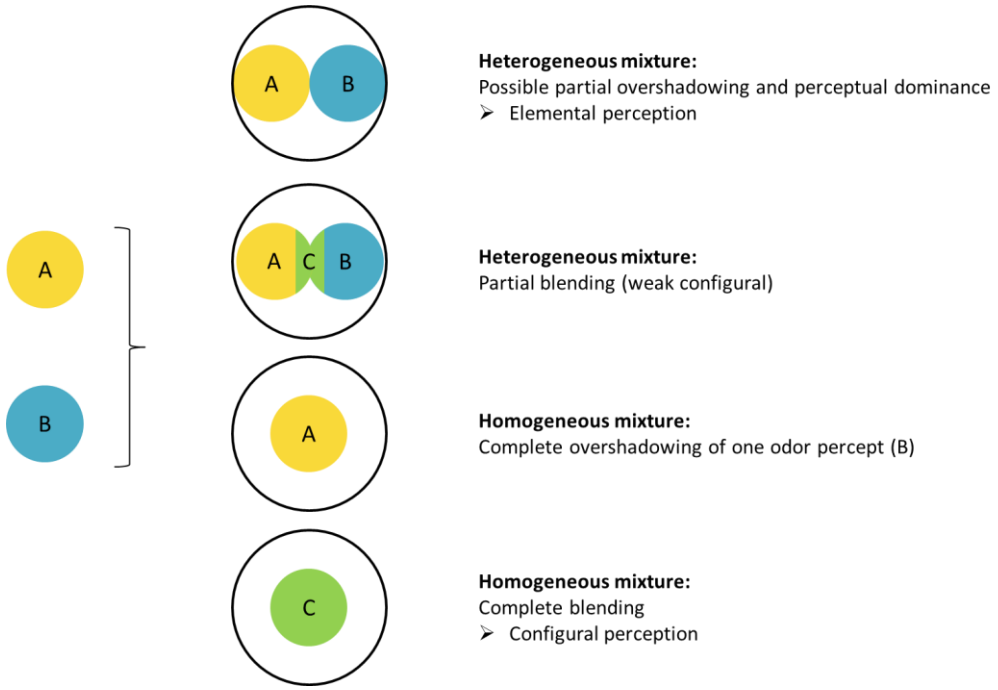


Figure 1.6: Perceptual interactions regarding the odor percept of a binary mixture. One odorant is described as the odor percept A and the other as the odor percept B, while the odor percept C is specific to the mixture (configural). Figure inspired by Thomas-Danguin et al. (2014).

b. Odor intensity

Perceptual interactions can also impact the perception of mixtures' intensity. Intensity interactions in binary mixtures are usually demonstrated by comparing a mixture's intensity with the intensities of its components alone or their sum (Cain, 1975; Patte & Laffort, 1979; Berglund & Olsson, 1993; Thomas-Danguin & Chastrette, 2002). Such interactions are categorized depending on whether the mixture quality is homogeneous or heterogeneous (Cain & Drexler, 1974; Berglund et al., 1976; Ferreira, 2012a).

On the one hand, heterogeneous binary mixtures, i.e. when both odorants keep their individual qualities, may smell as strong, more intense or less intense than the sum of both intensities due to respectively independence, synergy or antagonism effects. Such effects applied to one or both of the odorants. In Figure 1.7, odorant A illustrates independence, synergy and antagonism effects whereas odorant B illustrates only independence because its intensity stays equal to the unmixed intensity. Synergy and antagonism occur when the intensity of one odorant in the mixture is respectively higher or weaker than its intensity when unmixed.

On the other hand, as shown in Figure 1.7, homogeneous binary mixtures may smell 1) as strong as the sum of the perceived intensities of the unmixed components, i.e. complete addition; 2) more intense, i.e. hyper-addition, or 3) less intense, i.e. hypo-addition. Cain & Drexler (1974) subdivided hypo-addition in three classes: if the mixture quality intensity is greater than that of the single intensities, intermediate, or smaller, the terms "partial addition", "compromise", and "subtraction" are used.

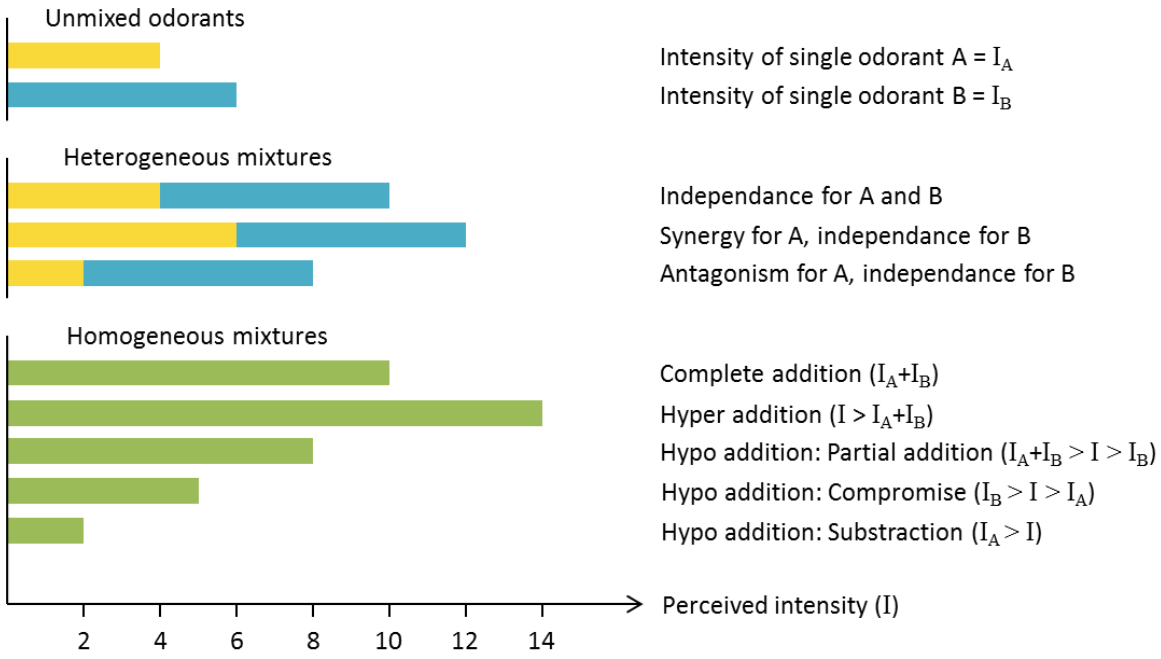


Figure 1.7: Perceptual interactions regarding the odor intensity of a binary mixture. In this example, the perceived intensity of odorant A smelled alone is 4 and the perceived intensity of odorant B smelled alone is 6. The perceived intensity of the binary mixture may result in a complete addition, hypo-addition or hyper-addition. Figure adapted from Cain & Drexler (1974) and Thomas-Danguin et al. (2014).

The studies presented above concern binary mixtures with odorants at supra-threshold intensities because compounds present above their odor threshold are usually considered as key compounds in odor construction. However, there are also interactions at intensities below threshold. Atanasova et al. (2005) showed the impact of sub- and peri-threshold odorants on the olfactory perception of mixtures.

Thus, considering complex food matrices composed of hundreds of odorants at various intensities (supra- and sub-threshold), interaction effects might be more complex. Moreover, the odor intensity of the mixture usually does not increase additively when increasing the number of components (Laffort & Dravnieks, 1982; Ferreira, 2012a).

When studying the perception of odorants' mixtures it is thus of importance to consider the odorants through their structural and perceptual variables when they are presented and perceived separately as well as when they are perceived in a mixture. In the following paragraphs, the techniques used to measure and evaluate food odor will be presented.

III. Food odor and its measurement

The odor of food is due to the processing by the olfactory system of many chemical molecules embedded in complex mixtures often recognized as single percepts due to odor blending. The odor of food products can be measured through chemical analysis to identify odorants in the products and/or through sensory analysis to give an overview of the product perceptual profile.

1. Chemical analyses

Food odor chemical analysis is classically performed by separating, identifying, and quantifying the volatile molecules included in food products using gas chromatography (GC) technique coupled with various detectors or electronic noses with various sensors or the human nose (GC-Olfactometry, GC-O).

a. Gas chromatography (GC) techniques

Because volatiles are contained within a food matrix, they first have to be extracted (Da Costa & Eri, 2009). Extraction techniques for GC analysis can be divided in solvent extraction methods, steam distillation methods, headspace techniques, and sorptive techniques. Among the wide variety of extraction methodologies, there is no universal extraction method to

produce a representative extract (Barba et al., 2017). In order to achieve the most complete volatile profile, i.e. to detect as many volatiles as possible, flavor chemists are then using combination of techniques. For example, Ferreira et al. (2002) combined micro extraction, solid phase extraction, and solid phase micro extraction techniques to identify 47 compounds in Grenache rosé wines.

Once volatiles are extracted from the food matrix, they are injected in a chromatograph along with a solvent (column and solvent choices are determined by the aim of the analysis and the characteristics of the extract). Each separated odorant, eluted by the GC, will be detected and identified with different methods according to the chosen detectors. The most common chemical detectors coupled with GC are mass spectrometer (MS, GC-MS for gas chromatography coupled to mass spectrometry) and flame ionization detector (FID). The visual output of such techniques is a chromatogram. In the case of an optimal separation, the different peaks on the chromatogram correspond to the volatiles included in the food product. Peak areas are linked to the relative amount of the volatiles in the extract but not to their odor intensities. As mentioned before, the link between concentration and perceived odor intensity is volatile-dependent.

Some volatiles identified might be odorless, whereas some volatiles might not be revealed by the chromatogram because they are present at such low concentrations that the chemical detectors cannot detect them because of the detector's limit of detection. Thus, a notable improvement in odor identification consisted in coupling GC-MS with olfactometric detection (GC-MS-O) (Delahunty et al., 2006). The GC-MS still allows the identification and quantification of volatiles molecules and the human nose is a second detector (Figure 1.8). Indeed, human assessors are a sensitive detector and can distinguish some odorants at concentration below ppt (Nagata, 2003). In addition to odorant presence, assessors are able to measure the duration of the odor activity (start to end), to describe the quality of the odor perceived, and to quantify its intensity (Delahunty et al., 2006; Bratolli et al., 2013).

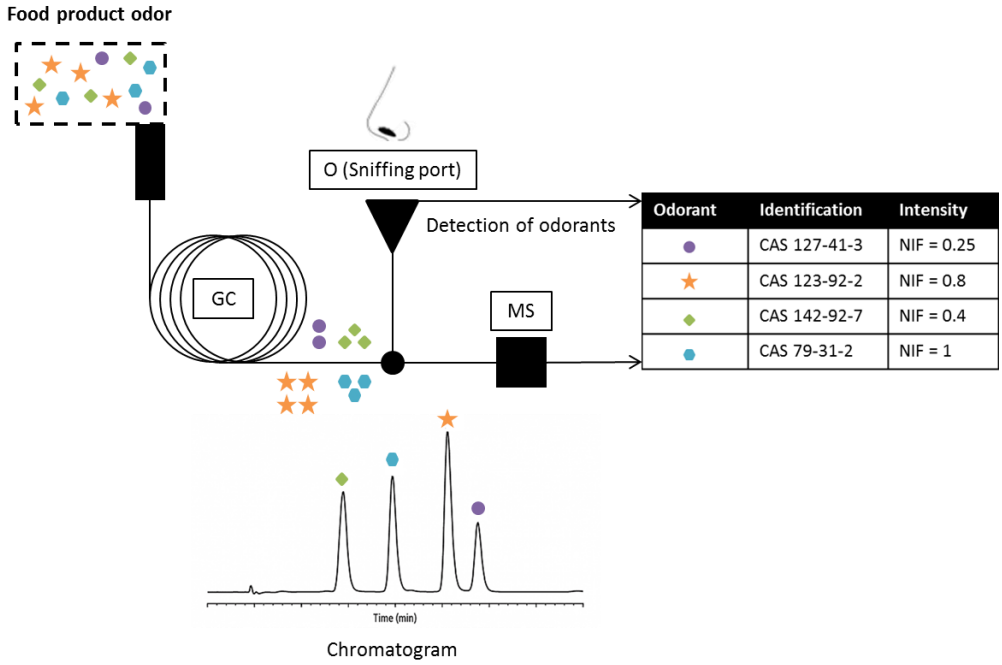


Figure 1.8: Food odor analysis through gas chromatography coupled to mass spectrometry and olfactory detections technique (GC-MS-O). Food product volatile compounds extract is injected in the gas chromatography (GC) column. Volatiles are separated and distributed both to the mass spectrometer (MS) and to the nose of a human assessor at the sniffing port. The MS detection allows the identification of volatiles whereas the olfactory detection allows the detection of the odorants among the volatiles. In this example, odorants are identified according to their CAS number. Intensity of odorants is calculated with their nasal impact frequency (NIF) scores which is the ratio of panelists who smelled the odorants on the total number of panelists.

Thus, GC methods associated to olfactometric detection have been developed in flavor research to determine the odor active compounds (i.e. odorants) among the volatile compounds present in food products and to determine their relative importance in a sample (van Ruth, 2001; d'Acampora Zellner et al., 2008; alcoholic beverages: Plutowska & Wardencki, 2008). These methods can be categorized into three groups: frequency detection, dilution to threshold, and direct intensity methods (Delahunty et al., 2006). In addition, we will present the concept of odor activity value (OAV).

- Frequency detection

The frequency detection method involves a panel of 6-12 people who analyze the same sample in order to provide the proportion of subjects from the panel who perceive an odor at a given retention time; this proportion is called the nasal impact frequency (NIF) value (Pollien et al., 1997; Le Fur et al., 2003). The NIF value is 1 when all the subjects perceived a given odor, and 0 when no subject perceived any odor at a given retention time (Figure 1.8). Simplicity is the main advantage of detection frequency-based methods which do not necessitate qualified evaluators. However, odorants present in different concentrations but all above the detection threshold will obtain a NIF value of 1.

- Dilution to threshold

Dilutions to threshold methods provide the odor potential of a given odorant based on the ratio between its concentration in the sample and its sensory threshold in air. These methods consist of preparing a dilution series of an extract, using twofold, threefold, fivefold or 10-fold dilution levels and then analyzing them with GC-O. The assessors state under which dilution the compound analyzed can still be perceived, and usually describe the type of smell. The most frequently reported dilution methods are Aroma Extract Dilution Analysis (AEDA) (Ullrich & Grosch, 1987) and Combined Hedonic Aroma Response Measurement (CharmAnalysis^{TM}) (Acree et al., 1984). Such methods allow ranking odorants according to their potency (i.e. the larger the dilution value, the greater the potential contribution of that compound to the overall aroma). A disadvantage of dilution methods is the length of time required to complete the analyses on each dilution for a single extract.

- Direct intensity

The odor intensity and its duration can be measured with direct intensity rating methods using different kinds of quantitative scales (category or unstructured line). These methods include a single time-averaged measurement registered after the elution of the odorant (posterior intensity evaluation method) or a dynamic measurement, where the appearance of an odor, its maximum intensity and decline are registered in a continuous manner. One potential drawback of direct-intensity methods is the substantial amount of training that assessors require in order to obtain individual reproducibility and agreement with one another.

- Odor activity value (OAV)

The OAV aims to estimate the importance of an odorant based on the ratio of the compound concentration in the food to its threshold concentration (Patton & Josephson, 1957). The ratio

indicates by how much the actual concentration of a compound exceeds its sensory threshold. Therefore this method does not explicitly rely on sensory evaluation within the analysis procedure. Whereas in AEDA or Charm analysis the threshold of each odorant is actually estimated during the analysis procedure, thresholds in the OAV approach are collected from the literature. Often defined in air or water, these non-specific thresholds are often misleading and not efficient (Audouin et al., 2001). Another drawback of the OAV approach is related to the linear concentration-intensity relationship assumption which is not an accurate estimation.

b. Electronic noses

The use of a human nose as a detector is useful as we just outlined but also limited by the fact that our sense of smell is subjective and gets tired easily. This led to the development of instruments that aim to mimic the human sense of smell such as electronic nose (Loufti et al., 2015). This instrument consists of an array of sensors for chemical detection and a pattern recognition unit usually coupled with artificial intelligence systems. The odor recognition process does not give information on sample composition but rather gives a digital signature or pattern (Haddad et al., 2010; Banerjee et al., 2016). Patterns comparison allows quality estimation (authenticity assessment), quality control (detection of bacteria and spoilage), and discrimination of various food products including wines (Rodríguez-Méndez et al., 2016). However difficulties in robustness, selectivity and reproducibility of the sensors, and the need for pattern recognition algorithms which can cope with the complex signal analysis are the main drawback of these instruments.

c. Limitations

Chemical analyses are at the core of flavor analysis. They are required to identify and quantify odorants in various food products. However, such methods consider the odorants separately and do not take into account the perceptual interactions among odorants which occur at the food product level. Furthermore, as mentioned previously, odorants present at sub-threshold impact the perception of food products. Moreover, the food matrix composition can impact on the perception of odorants in food products (Ickes & Cadwallader, 2017). Though analytical methods are useful in understanding the odor construction of food products, key information might be missed. Thus, researchers now consider AEDA, Charm and OAV as methodologies to determine which odorants most likely make a contribution to the food odor recognizing that additional sensory work is needed to determine which aroma

compound are truly contributory (Audouin et al., 2001). This additional sensory work mostly relies on recombination studies that will be presented below.

2. Sensory analyses

Food odor sensory analyses rely on measuring characteristics of a food product as they are perceived by the sense of smell. Besides hedonic evaluations, several methods are used according to the aim of the analysis: food odor comparison or description. These methods are conducted with human subjects selected on their skills to obtain reliable and valid results.

a. Comparative methods

Comparative methods include discrimination, similarity and preference measurements between food odor samples.

Discriminative tasks aim to identify if several samples are perceived differently. Results are obtained through paired-comparison, duo-trio, triangle, or dual-standard tests (Stone & Sidel, 2004). In addition to these tests, the magnitude of the difference can be determined between the products. To do so, subjects first make a discrimination decision followed by a measure of the perceived magnitude on a category-type scale (e.g. weak, moderate, and strong) or a numeric-scale. Another modification to the discrimination test may require subjects to indicate why they discriminate the products.

Similarity tasks follow the same procedure as the discriminative tasks, the aim being to identify the perceived similarity among food products.

b. Descriptive methods

Descriptive tasks aim to provide a sensory description of products and thus allow constructing sensory profile following qualitative and/or quantitative analysis (Stone & Sidel, 2004). There are several descriptive sensory analysis methods available, such as quantitative descriptive analysis (QDA), Spectrum, or Flavour Profile (Lawless & Heymann, 2010). Accuracy of the methods depends on whether panelists are trained or not. The more trained the panelists, the less variable are the results.

Whereas Flash Profile and Free Choice Profiling require panelists with none or little training, QDA methodology requires trained panelists. The first step is the training of panelists to

select the sensory attributes further used to describe the food products. Usually the training is done with food products close to the ones to be evaluated. During the training sessions, the list of attributes might be reduced as the subjects evaluate and discuss their responses and realize that many of the words have a common sensory basis. There is also a step of training using references dedicated to each attribute. Once the sensory attributes are consensual and relevant for the panel, the samples are evaluated on different type of scales depending on the method. As a result, QDA methodology provides the quantitative description of food products for all the sensory attributes selected (Figure 1.9). Thus, QDA methodologies are widely used in the food industry to gain insights on how products are perceived by the consumers relative to their sensory differences. Furthermore, such analyses are useful to compare products according to their change over time (storage testing) or in product development to determine whether experimental formulations match a target.

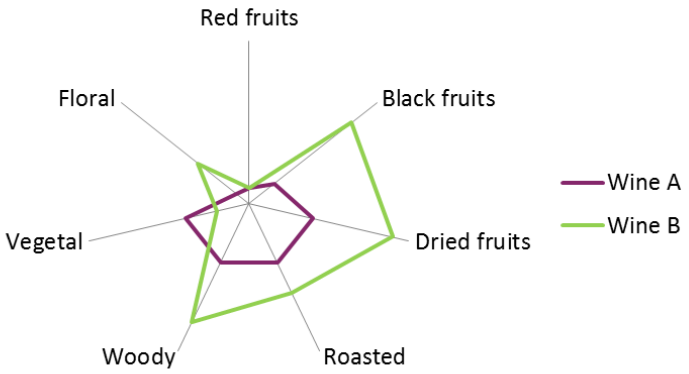


Figure 1.9: Visual display of the odor sensory attributes of two wines A and B based on the results of a QDA test. For each attribute, the relative intensity increases as it moves farther away from the center point. Figure adapted from Ferreira et al. (2016).

c. Limitatios

The choice of a method among the diversity of sensory analysis is determined by the objective of the analysis. Hence, if the question is "are there differences among the products", a discrimination test is indicated. If the question is to know which attributes have changed in the sensory attributes of a new product, a descriptive analysis procedure is required. Such

procedure has proven to be the most comprehensive and informative sensory evaluation tool. Nevertheless, such analysis is time-consuming because training of panelists is often required.

3. Relationships between analytical and sensory analyses

Several studies had been conducted in correlating attributes from sensory analyses with volatile data obtained by GC analyses. Such attempts aimed to predict the odor of food product based on their chemical composition to skip sensory methods which are expensive to implement, time consuming and difficult to implement on-line for immediate feedback (Chambers & Kopple, 2013).

a. Statistical approaches

The most common approaches to correlate analytical and sensory spaces are statistical approaches. Various statistical methods can be used and include principal component analysis (PCA), generalized procrustes analysis (GPA), and partial least squares regression (PLSR). These three methods have been widely used to correlate sensory profiles of wines according to their odorants composition (PCA: de la Presa Owens et al., 1998; GPA: Le Fur et al., 2003; PLSR: Lee & Noble, 2003, Aznar et al., 2003, Campo et al., 2005 and González Alvarez et al., 2011). PCA is used for dimensionality reduction of multivariate datasets based on linear combinations of the original variables (GC peaks or sensory attributes) which are principal components. PCA can be used to identify combinations or patterns of variables having the largest contribution to variability in the data set. GPA is a method, which can be used to find a common structure between two datasets like sensory and volatiles data. As a result, each sample is mapped with three markers: based on sensory results, based on GC results and the consensus spot of sensory and GC results. The resulting map gives an indication of how the samples are grouped and also shows the correlations between patterns in sensory data and GC-O results. PLSR, the main statistical tool in correlating analytical and sensory data (Seisonen et al., 2016), is a method for relating two data matrices in order to estimate linear combinations of one set of variables (e.g. analytical data) that predict much of the variation in another set of variables (e.g. sensory attribute ratings). An illustrative example of PLSR results is shown in Figure 1.10.

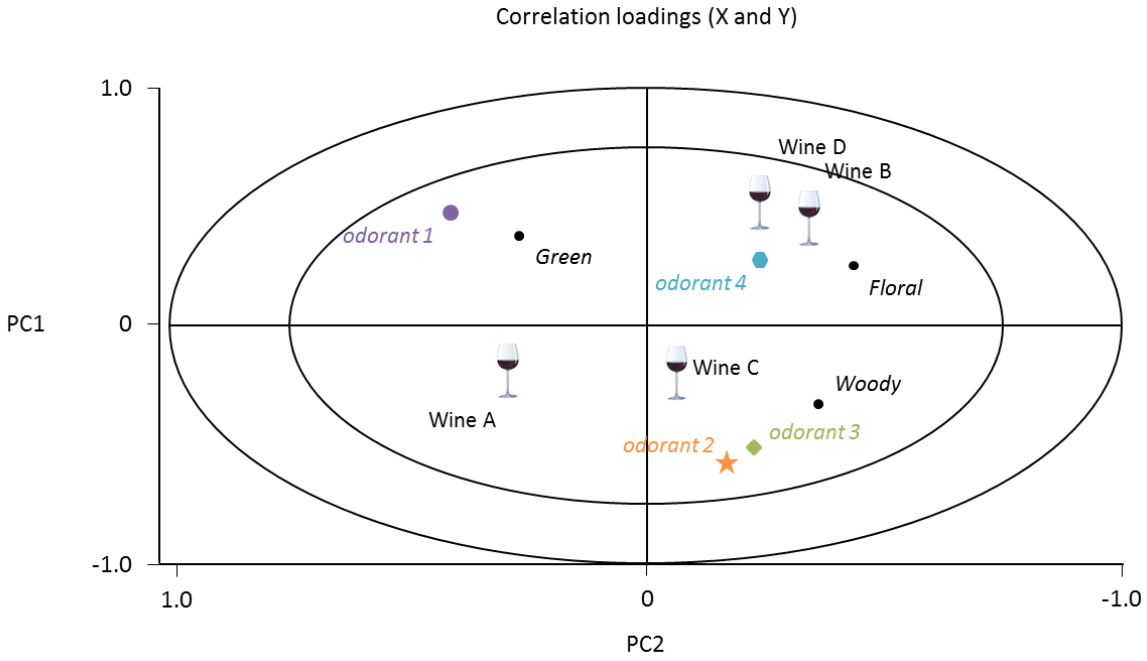


Figure 1.10: Illustrative example of a partial least square regression (PLSR) to correlate analytical and sensory data. The simplified figure presents four wine samples described according to four odorants and three sensory attributes and gives an overview of the variation found in the mean data from a plot of PLSR correlation loadings for odor-active compounds (X-matrix) and samples and sensory attributes (Y-matrix). Results highlight correlations between samples, odorants and sensory attributes. For example, odorants 2 and 3 have a strong positive correlation with the sensory attributes Woody and characterize the wine C. Figure inspired by Lee & Noble, 2003.

The main disadvantage of the statistical methods is that they only propose linear-based statistical relationships among the variables. Such relationships can help establishing correlations between odorants and sensory attributes to indicate that the variables are changing in the same manner rather than identify causality. Indeed, statistical analyses are not suitable to give insights on construction of odor perception. In addition, such methods are focusing on the impact of single odorants on sensory attributes perception. However, in complex products, combinations of odorants may lead to different odors than those expected from individual components because of non-linear perceptual interactions detailed in the previous part of this literature review.

To overcome the statistical relationships, alternative procedures are developed but they require pre-treatment of the data. The use of multivariate methods will give more complex relations. The reliability of the results should still be interpreted with caution (Macfie & Hedderley, 1993).

b. Recombination and omission approaches

Whereas statistical approaches consider single odorants, recombination and omission approaches aim to identify key odorants which may impact the odor/aroma of food once in mixtures.

First, the recombination strategy step is to recreate the full odor of a food product as a model. Usually, only odorants with an OAV higher than 1 are kept. The model is then compared to the real food odor (Lorrain et al., 2006). If no differences are perceived, the omission study follows. The preparation of aroma models is simple for liquid foods, as it is easy to obtain a homogenous blend of odorants (Grosch, 2001). Difficulties arise, however, in the case of solid foods, because in general it is not easy or even possible to reproduce the composition and distribution of the non-volatile components of the real food matrix in a model system.

Omission studies evaluate the aroma model with the model minus a single odorant or a group of odorants. If a sensory difference is perceived between the models, the odorant or the group of odorants omitted is considered responsible for the sensory difference perceived (baijiu: Zheng et al., 2016; wine: Lytra et al., 2013; Ferreira et al., 2016). For example, in omission tests carried out on Grenache rosé wines, 3-mercapto-1-hexanol (CAS 51755-83-0) can be considered as an impact compound because its omission changes greatly the aroma of the wine by reducing the citric and fruity notes and increasing the floral and caramel notes (Ferreira et al., 2002). Several recent studies also highlighted the impact of higher alcohol compounds on the perception of animal, fruity and woody notes in red wines (Cameleyre et al., 2015; de-la-Fuente-Blanco et al., 2016; de-la-Fuente-Blanco et al., 2017).

These approaches have proven the indirect contribution of odorants to the overall odor of food product. However, reconstitutions and omission experiments are long processes and require panelists for the sensory evaluation of the various aroma models and sub-models. Furthermore, as already discussed, the use of odorants with high OAV into the model formulation may not be accurate. Indeed, odorants with lower OAV can impact the overall property of a food product (Escudero et al., 2004).

Most attempts to relate sensory profile to analytical measurements probably failed because of seeking for direct relationship between two datasets that are fundamentally different in nature (Chambers & Kopple, 2013). Recombination strategies have the advantage to focus on odorants' mixtures but may miss information because of considering mostly odorants with high OAV. In order to correlate sensory and instrumental datasets, it is thus important to use non-linear modelling approaches that might result in efficient predictive tool. The fact that odorants are contained in various food matrices should also not be forgotten. Indeed the perception of odorants in different matrices varies, which is rarely taken into account when relating compounds with sensory attributes.

IV. Predicting odorant characteristics

While classical statistical approaches have been widely used in food flavor and sensory related studies, these techniques may be inadequate in fully describing a complex and potentially non-linear system found in the odor of food products (Yu et al., 2018). Thus, olfactory research is still challenged by predicting odor characteristics of molecules such as their detection threshold, pleasantness or their smell. Apart from the statistical approaches, a dynamic research field is the development of artificial intelligence-based approaches. Several methodological tools proposed by this area of study have been implemented in the food science domain, such as machine learning (including optimization and algorithms). A brief look at the literature indicates that most studies focused on predicting characteristics of single odorants (Table 1.2). Furthermore, the table highlights the predominant use of molecular structure approaches to predict odorant characteristics. These approaches are combined with different modelling strategies and aimed to classify odorants according to their characteristics (intensity, quality, pleasantness) or to predict their characteristics.

Table 1.2: Predicting odorant characteristics. A non-exhaustive list of studies is presented in order of occurrence with first studies applied on single odorants, then studies applied to mixtures of less than 10 odorants and finally mixtures of more than 10 odorants. The column "Input" concerns the data used for the prediction of the characteristics in the next column. The term "odor quality" refers to the description of the odor percepts in verbal terms. In the column "Method", the data available about the method used are provided.

Single odorant or mixture	Input	Characteristic to predict	Aim	Method	References
Single odorant	Structural parameters and sensory analysis	Trigeminality	Prediction	Multiple regression equation	Doty et al., 1978
Single odorant	Structural parameters and analytical analysis	Intensity	Prediction	Multiple regression equation	Edwards & Jurs, 1989
Single odorant	Structural parameters and sensory analysis	Odor quality (musk)	Classification and prediction	Artificial neural network (2 layers with respectively 6 and 3 neurons, back-propagation algorithm)	Chastrette & de Saint Laumer, 1991
Single odorant	Structural parameters	Odor quality (ambergris, bitter almond…)	Prediction	Equations and structural rules for specific odor quality	Rossiter, 1996
Single odorant	Structural parameters and sensory analysis	Pleasantness	Prediction	Correlation (principal components analysis)	Khan et al., 2007
Single odorant	Analytical (electronic nose) and sensory analyses	Pleasantness	Prediction	Artificial neural network (3 layers, 5 internal neurons, 20 epochs, 20 repetitions, back-propagation algorithm)	Haddad et al., 2010
Single odorant	Structural parameters and sensory analysis (from literature)	Pleasantness	Prediction	Correlation	Kermen et al., 2011
Single odorant	Structural parameters and sensory analysis (from literature)	Pleasantness	Prediction	Multiple regression equations and structural rules	Zarzo, 2011

Single odorant	Structural parameters and sensory analysis (from literature)	Odor detection threshold	Prediction	Multiple regression equation	Abraham et al., 2012
Single odorant	Databases and structural parameters	Odor quality	Classification	Random forest classifier (with and without feature selection)	Kumar et al., 2015
Single odorant	Structural parameters and sensory analysis	Intensity, pleasantness, odor quality	Prediction	Correlation	Keller & Vosshall, 2016
Single odorant	Structural parameters and sensory analysis	Intensity, pleasantness, odor quality	Prediction	Different models using machine-learning algorithms: regularized linear model to random forest algorithm	Keller et al., 2017
Single odorant	Structural parameters and odor description (from databases)	Odor quality	Prediction	Machine-learning approaches: support vector machine, random forest and extreme learning machine (different feature selection tested)	Shang et al., 2017
Single odorant	Structural parameters and sensory analysis (both from databases, see Keller et al., 2017)	Odor quality	Prediction	Random forest algorithm (100 trees, with and without feature selection)	Li et al., 2018
Single odorant	Structural parameters and sensory analysis (from literature)	Odor detection threshold	Prediction	Algorithm to select model and descriptors (genetic function approximation) and partial least square regression	Ojha & Roy, 2018
Single odorant	Analytical and sensory analyses (both from databases)	Odor quality (presence or absence of odor descriptors)	Prediction	Artificial neural network (6 layers, back-propagation algorithm) coupled with language modeling method	Nozaki & Nakamoto, 2018
< 10 components	Sensory analysis	Pleasantness	Prediction	Equation	Lapid et al., 2008

< 10 components	Sensory analysis	Odor intensity	Prediction	Equation	Yan et al., 2015
< 10 components	Sensory analysis	Odor intensity	Prediction	Equation	Laffort & Dravnieks, 1982
< 10 components	Sensory analysis (from literature)	Intensity	Prediction	Equation	Thomas-Danguin & Chastrette, 2002
< 10 components	Analytical (electronic nose) and sensory analyses	Odor intensity	Prediction	Fuzzy-sets theory (8 input variables, i.e. sensor signals, center of gravity for the defuzzification)	Szulczyński et al., 2018
> 10 components	Analytical and sensory analyses	Flavor intensity	Prediction	Artificial neural network (back-propagation algorithm, principal component regression sensory scores as input)	Boccorh & Paterson, 2002
> 10 components	Analytical (electronic nose and GC-O) and sensory analyses	Aroma	Prediction	Artificial neural network	Michishita et al., 2010
> 10 components	Structural parameters and sensory analysis	Similarity	Prediction	Equation (angle distance model)	Snitz et al., 2013
> 10 components	Analytical (electronic nose and GC-O) and sensory analyses	Red fruit aroma	Underlying relationships	Fuzzy-sets theory (5 input variables, i.e. groups of molecules)	Tomasino & Tomasino, 2018

1. Molecular structure approaches

Early approaches based on the structure-odor-relationships (SOR) (Rossiter, 1996), identified many specific rules linking odorants' chemical structure to odor (e.g. which structure features may provide a "woody" note), but failed to produce a general framework for measuring smell. Indeed, molecules that have a close structure can elicit very different responses, for example (−) carvone (CAS 6485-40-1) smells minty, whereas its enantiomer (+) carvone (CAS 2244-16-8) smells like caraway (Friedman & Miller, 1971) (Figure 1.11). On the opposite, molecules with different structures such as muscone (CAS 541-91-3), musk ketone (CAS 81-14-1), traseolide (CAS 68140-48-7), and helvetolide (CAS 141773-73-1) have similar musk odors (Figure 1.11). Such results were confirmed by studies at the peripheral level of the olfactory system where odorants with similar structure did not activate the same pattern of olfactory receptors (Araneda et al., 2000).

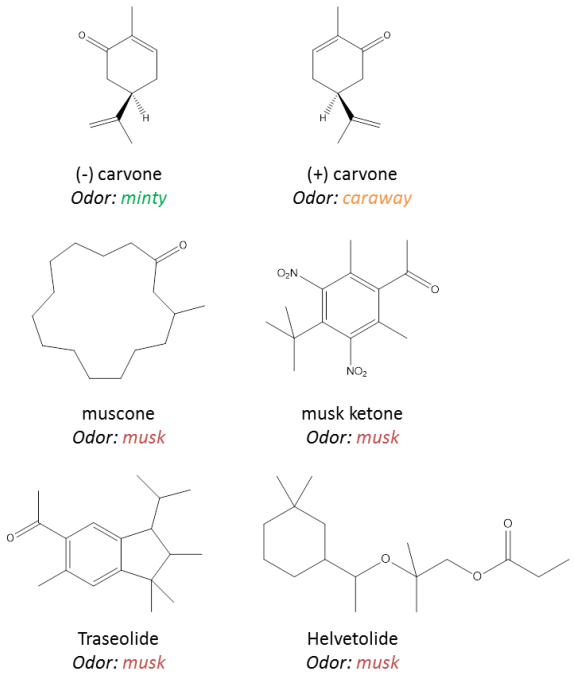


Figure 1.11: Structurally similar molecules but dissimilar odors and vice versa. Figure adapted from Sell (2006).

For this reason, modern attempts to characterize molecular structure have focused on modelling approaches, hoping that a systematic analysis will capture what direct correlation

or statistical analyses missed. In addition, recent progress in software that provides molecular descriptors (e.g. Dragon software, Talete, Milan, Italy) now allows the consideration of thousands of structural parameters for any molecule. These quantitative molecular descriptors represent measures such as constitutional indices (molecular weight, number of atoms), ring descriptors (aromatic ratio), functional group counts (number of thiols), 3D atom pairs (sum of geometrical distances between S-Br)…

Researchers from several different groups have been working to find models based on the structural parameters of molecules to predict the intensity, pleasantness or smell quality of a new set of molecules (Table 1.2). Abraham et al. (2012) worked on predicting the odor detection thresholds of molecules using several structural parameters (molar refractivity, solute dipolarity, number of hydrogen bond, gas to hexadecane partition coefficient, functional group). Their results showed a good correlation between the odor detection thresholds calculated and observed ($R^2 = 0.759$). Another study (Kumar et al., 2015) showed that it is possible to use selective structural parameters to predict the perceptual qualities of odorants (e.g. "butter", "fruity", "leather"). To do so, they collected the structural parameters and odor descriptors of more than 3000 molecules. They found a significant overlap between the spatial positioning of the clustered odorants in the structural and perceptual spaces. Then, they developed a random forest classifier to predict the odor quality of a novel odorant using its structural parameters. A step forward was working on the link between the structural parameters of molecules and their hedonic dimension (i.e. pleasantness). Khan et al. (2007) found that their model can provide a good prediction of the perceived pleasantness or un-pleasantness of novel molecules ($R = 0.55$). In addition Kermen et al. (2011) showed that the more structurally complex a molecule, the more numerous the olfactory notes it evokes and the less pleasant the molecule is perceived. Other strategies are also using machine learning algorithms to predict several odor characteristics of molecules based on their structural parameters. Keller et al. (2017) presented modelling results from 22 teams which were given a large dataset containing sensory and structural data (Keller & Vosshall, 2016). Sensory data were gathered from 49 subjects who profiled 476 molecules on 20 odor descriptors. Structural data represented 4884 structural parameters for each of the molecule smelled by the subjects. The resulting models accurately predicted odor intensity, pleasantness and also successfully predicted 8 semantic descriptors ("garlic," "fish," "sweet," "fruit," "burnt," "spices," "flower," and "sour"). For instance, one of the participating team (Li et al., 2017) developed a

random forest model consisting of multiple decision trees which successfully predicted personalized odor descriptors of structurally diverse molecules.

Although these predictive approaches were successful, they all applied to single odorants. Snitz et al. (2013) began to work on the odor characteristics of mixtures. They developed an angle distance model coupled with an algorithm that can look at the structural parameters of two novel odorants' mixtures in order to predict their perceptual similarity. The main idea was to consider the odorants in mixtures as single structural vectors to take into account the mixture as a whole and not only a sum of its constituents' chemical features. This is consistent with a configural rather than an analytical processing of odor mixtures (Thomas-Danguin et al., 2014). Their algorithm provided consistent correlations between predicted and actual perceptual odor similarity between mixtures ($r = 0.49$, $p < 0.001$). An optimized version of this model (selection of 21 structural parameters among the 1433 used in the first model) yielded a correlation of $r = 0.85$ ($p < 0.001$) between predicted and actual mixtures' similarity. Despite this improvement, two main points should also be taken into account. The angle distance model was constructed with mixtures of odorants at iso-intensity and the output of the model is the similarity among two mixtures but no prediction in terms of odor percept. On the one hand, iso-intense mixtures, when all the odorants embedded in the mixture are perceived at the same intensity when evaluated alone, are experimentally created but do not occur in nature. The following step is then to inject intensity parameters into the predictive approach (Snitz, 2016). On the other hand, the predictive model helps to group mixtures of odorants according to their similarity but is unsuitable to predict characteristics of mixtures such as the odor percept they evoke.

2. Modelling strategies

Several modelling approaches were listed in Table 1.2. Such models are usually divided in three groups (white-box, black-box, and grey-box) regarding the level of details required (Perrot et al., 2011).

a. White-box models

When the modelling approach is primarily guided by the knowledge of the underlying mechanisms, the resulting model is usually termed as white-box. Such models can be based on equations (e.g. kinetic, power law). For example, these analytical models were used in

predicting intensity of single odorants or mixtures according to the odorants' concentration; they rely on mathematical non-linear relationships (Patte & Laffort, 1979; Laffort & Dravnieks, 1982; Berlung & Olsson, 1993a; Berlung & Olsson, 1993b; Chastrette et al., 1998; Thomas-Danguin & Chastrette, 2002).

b. Black-box models

In contrast, empirical data-driven or black-box models describe observed tendencies in experimental data by arbitrary mathematical functions such as artificial neural networks (ANN) or random forests. However, the internal working of the device is not described, and the model simply solves a numerical problem without reference to any underlying physical or biological processes.

ANN is a method that tries to simulate the way a human brain works. The model consists of different layers of neurons: an input layer, one to several hidden layers and an output layer, and there are connections between neurons in each layer. Each neuron is in fact a mathematical function linking inputs variables to outputs variables. When dealing with sensory-instrumental relationships, the input layer could be considered as independent variables (e.g. volatile composition) and the output layer as dependent variables (e.g. sensory attributes). Therefore, a model with a sufficient number of hidden layers can be created and transformed by different mathematical algorithms to predict sensory characteristics from volatiles data, by using a training dataset. After this, the reliability of the model with new known data can be tested and used for predicting sensory parameters of unknown samples. The main advantage of ANN over statistical analysis such as PLSR is the capability to account for non-linear relationships. Quick and easy-to-use when sufficient experimental data is available, such models nevertheless encounter important limitations when applied to food systems: risk of over-parameterization, interpretation difficulty, lack of generalization ability (Perrot et al., 2011). Some parameters are not easy to set up and/or optimize like the number of neurons and number of hidden layers. Moreover, the numbers of required measurements increase exponentially with the number of studied factors and thus with the complexity of the problem to solve. Because sensory analyses are often time consuming, the sensory data gathered is quite small. This can explain why the use of ANN for modelling sensory-instrumental relations is not very common (Chastrette & de Saint Laumer, 1991; Boccorh & Paterson, 2002; Michishita et al., 2010; Cancilla et al., 2014).

c. Grey-box models

An intermediate approach consists in designing a model based on black-box structure and to complete missing information by empirical relationships derived directly from experimental data or expertise knowledge. Such models are sometimes called grey-box because methods from the white-box and black-box are combined. The advantage being that some elements within the model can be approximated by rules. Inference rules can be created as fuzzy rules.

Fuzzy rules are used within fuzzy logic systems to infer an output based on input variables. Fuzzy logic is an extension of the classical logic and introduces additional values between standard 0 and 1 and is often used in sensory science for pattern recognition and clustering with limited use in predictive modelling (Yu et al., 2018). For instance, fuzzy logic can be used along with ANN in classification tasks (Yea et al., 1994; Scott et al., 2006). Combined with sensing devices such as electronic noses, the use of fuzzy logic as a classifier led to good discrimination among odors (Upadhyay et al. 2017). Indeed, fuzzy concepts introduced by Zadeh (1965) provide interesting alternative solutions to the classification problems within the context of imprecise categories. A recent study (Szulczyński et al., 2018) applied fuzzy logic to determine the odor intensity of odorants mixtures. Analytical measurements were performed with an electronic nose. From the sensor results, two systems were built: a multi-linear regression and fuzzy logic system. The results obtained using fuzzy logic were closer to the results from sensory evaluation compared to results obtained with the multi-linear regression. However, the authors highlighted that the creation of fuzzy rules is more complicated than the determination of linear model parameters because it requires experts' knowledge.

Owing to the complex relationships between mixtures' composition and sensory attributes, most of the predictive approaches presented here applied to single odorants or to unrealistic mixtures. The increasing use of machine learning techniques along with grey-box models might help to conduct predictive work on a wider range of odorants' mixtures. To do so, expertise might be embedded in modelling strategies (Szulczyński et al., 2018).

V. Expertise

Expertise about food odor perception might provide insights in food flavor analysis and be useful in approaches aiming to predict the odor of a food product. Although numerous studies analyze expertise both in psychology and in artificial intelligence, it is not possible to provide a consensual definition of this notion (Shanteau, 1992). Nevertheless, the prevalent criteria to distinguish novices and experts are based on seniority, i.e. duration of practice, skills, and peer recognition (Cellier et al., 1997; Sicard et al., 2011).

1. Food experts

Food experts can be divided in two groups. On the one hand, experts dealing with dynamic environments and managing food processes. On the other hand, experts dealing with static environment to provide sensory description of food products (wine, coffee, or beer for example).

a. Process

Dynamic environments are defined by the stage of the process that changes, irrespective of operator action (Cellier et al., 1997). Hence, the complexity of such processes is a result of temporal constraints that compel experts to make decisions before a full diagnosis can be made. They must not only take the present situation into account but past situations and forecasted future situations as well. For example, in cheese ripening process, the major difficulty is the delay between an action and its observed effect.

Experts have higher performance levels than novices in managing dynamic processes (Chi et al., 1981) because they have a more global and functional view of the situation. They can control food processes based on their skills and mental heuristics (i.e. interpretation of the situation in order to respond directly). Skills and heuristic rules are acquired during the experts practice.

b. Products

Overall, it seems that experts are better than novices to describe, memorize, and discriminate between sensory stimuli, but the difference in performance is not always as impressive. On

the one hand, some studies concluded that regarding flavors, experts only have a limited advantage over novices and highlighted sensory expertise as domain-specific expertise (coffee and wine: Croijmans & Majid, 2016). On the other hand, experts were shown to have a better ability to identify odors or for instance to separate wines according to their grape varieties (wine: Ballester et al., 2008; Tempere et al., 2016).

The wine sector is highly dependent on experts. Indeed, purchase decisions for wine consumers are influenced by wine experts who help to remove some of the perceived risk involved in purchasing wine by providing guidance on quality, taste profile, and relative value to consumers (Hayes & Pickering, 2012). This professional expertise relies both on chemosensory and technical knowledge. The chemosensory knowledge is acquired through repeated wine tastings and sniffings. The technical knowledge encompasses knowledge in chemistry, viticulture, oenology, or wine making processes. The reason why wine experts perform better than novices may rely on two explanations. First, experts might be using a more efficient wine-tasting procedure including a more analytical approach than novices. This more efficient procedure could lead to a superior ability to discriminate and identified wines (Ballester et al., 2008). Second, experts' perception could be enhanced by top down processes in which knowledge on different wine styles and varieties affects their sensory assessment. Experts having a deeper knowledge of wine styles than novices would then be able to focus on the individual features which differentiate samples the best (Solomon 1997; Hughson & Boakes, 2002).

2. Flavor experts

Whereas food experts are considered domain-specific, flavor experts possess a general ability to express common odors in language contrary to novices (Sezille et al., 2014). Indeed, the ability to communicate about flavors is a matter not only of perceptual training, but specific linguistic training too. Flavor experts have an extensive knowledge of the numerous molecules and ingredients that can be combined to develop or create flavors. A typical part of their studies consist in learning both semantically and perceptually an extensive list of flavor molecules along with their corresponding smell and/or taste and their usage in flavor formulation.

Flavor creation is the fusion of science, experience, and artistry, as there are likely numerous possibilities for combining molecules to achieve a desired flavor perception. The two main methods in flavor creation are linear and block formulation. The choice of the method depends on the flavorist's seniority and the aim of the formulation (e.g. creating a new flavor or improving an existing one). In linear addition, compounds are added one by one and the impact of each molecule on the mixture is evaluated. In block addition, mixtures of compounds (i.e. blocks) are first formulated to target a specific odor quality and are further assembled. For example, a coffee flavor might be the result of the combination of burnt, caramel, cocoa, fruity/floral and toasty blocks (Lissarrague, personal communication). Final results of each method are subjected to rearrangement to obtain a balanced flavor. Hence, flavorists are mostly relying on the creation of blocks, also called odor qualities. The way flavorists are combining odor qualities or molecules to create a target odor represents their expertise.

Regarding sensory evaluations, the mental process used by flavorists is very different from that used by sensory experts. Flavorists use an analytical process to deconstruct their flavor perception, in the sense that they try to find which molecules or ingredients were combined to generate each flavor direction. The sensory experts would for instance describe the differences between a targeted flavor and a created match proposal in order to determine what aspects of the targeted flavor profile have already been reached, but do not provide ingredient information. This will lead flavorists to focus on very specific flavor aspects of the products, which might not be a focus for sensory experts as they rather consider the food product's flavor as a whole (Veinand, 2015).

3. Modelling expertise

Modelling human expertise was made possible by the development of tools able to take into account non-numerical data.

Ontologies allow to structure and organize expertise into a meaningful structure at the knowledge level. The objective is to collect and formalize scientific knowledge from experts or bibliography to construct an informative system. An example is the construction of a Process and Observation Ontology (PO^2) applied to dairy gels (Ibanescu et al., 2016). The aim of this ontology was to provide a consensual structure representative of the production and

transformation of dairy gels and to solve the lack of communication between domain experts because data were gathered for different purposes by different experts with their own experimental itineraries, vocabularies, and methods. Thus, nutrition, microbiology, biochemistry, physico-chemistry, chemistry, process engineering, food science, and sensory analysis experts participated to the construction of PO². Explicit and implicit knowledge from these domain experts helped to design decision support systems allowing to compare different production scenarios and therefore suggesting improvements concerning the product quality (e.g. its sensory properties) while reducing the environmental impact.

Fuzzy logic, presented earlier, is a convenient mathematical approach to cope with applications where expertise is present (Perrot et al., 2006). This theory is particularly well adapted for dealing with the symbolic data manipulated by experts (Perrot et al., 2011). Thus, fuzzy logic has been used to deal with the implementation of expert knowledge encoded in fuzzy rules using several modelling strategies (Linko, 1998). Several applications are reported in the modelling of expertise on dynamic food processes to control food quality (review: Perrot et al., 2006 / Allais et al., 2007). Examples cover a wide range of food products: cheese ripening (Perrot et al., 2004; Sicard et al., 2011), biscuit baking (Perrot et al., 1996), milled rice (Zareiforoush et al., 2015). Ioannou et al. (2002) developed an approach based on fuzzy set theory to predict the sensory properties of crusting sausages. By means of a camera and an adapted image processing, the color of sausages was estimated. Operators' expertise was then integrated in the model to estimate the degree of sausage crusting according to the sausages' color. Nevertheless, the bottleneck of these approaches is the difficulty to capture the dynamics of the system using the expert knowledge.

Fuzzy logic has the advantage to be able to mimic human reasoning and thus allow the creation of fuzzy rules which can be implemented in grey-box models. However, within the food science framework, expertise modelling was mainly applied to processes. To the best of our knowledge, studies combining fuzzy logic and expertise on food odor or odor mixtures do not exist or have not been published. A recent paper presented during the 254th American Chemical Society National Meeting (Tomasino & Tomasino, 2017) applied fuzzy-set logic analysis to establish the relationships between chemical composition and sensory perception of wines. Starting from the chemical composition of wines they aimed to determine the molecules or group of molecules (norisoprenoids, furaneols, lactones, esters) responsible for red fruit aroma in Pinot noir wines.

CRITICAL CONCLUSIONS AND THESIS PROBLEMATIC

Among the sensory dimensions involved in food flavor perception, the odor component is critical because it determines most of the time the identity and the typicality of the food, which drives its overall quality and recognition by consumers.

Nowadays, aroma analysis, namely the chemical analysis of the odor component of food, is performed by separating, identifying, and quantifying the molecules included in an extract using a variety of efficient GC methodologies. This well-established analytical procedure provides a list of odorants, but does not give any information about the perception of these odorants within a mixture, still critical to the overall food odor construction. Indeed, peripheral and perceptual interactions occur during the perception of odorants' mixtures by the olfactory system. Mixtures are often recognized as single percepts due to odor blending resulting from configural processing. At the same time, we are also able to discriminate odors within complex mixtures through their elemental processing but not all components are perceived because of perceptual masking effects. Since the biological and neurobiological mechanisms of the underlying integration processes are still poorly known, it is very difficult to predict the odor of a given food product on the basis of its molecular composition.

Indeed, the search for a molecule-based classification system to define perceptual space and facilitate objective communication about odors is widely recognized as, and remains, a major challenge in olfaction research. We have seen that different teams have been working on establishing predictive approaches from the structural parameters of molecules but many of them apply to single odorants or unrealistic mixtures. These approaches relied on several modelling strategies. To date, the most advanced approaches are based on correlations or ANN methods. If such modelling strategies were found to have rather good abilities to predict the odor qualities of a given odorant using semantic descriptors, no studies attempted to develop such predictive method for odor mixtures. Moreover, these black-box models are not able to give cues on the underlying mechanisms of odor percept construction.

In order to get insight into complex odorants' mixtures perception and to develop efficient odor percept prediction tools, an interesting path would be to use grey-box models, which may integrate expertise knowledge. In the case of odor perception, we have seen that flavorists have extensive experience in combining molecules to create a targeted odor. This expertise can be useful to predict the odor of odorants' mixture. Therefore, it is likely that general expertise would be of interest to provide global knowledge on odor mixture

perception and especially critical molecules association rules. Such an approach would require the integration of heterogeneous data (e.g. chemical, sensory, expert knowledge) in a single model. Mathematical and computational tools are available within the applied machine learning and knowledge integration fields to support such heterogeneous data modelling. However, to our knowledge no study developed an approach to predict how a multi-molecular mixture will smell and could be qualified using verbal descriptors.

These scientific questions are at the core of this thesis, which attempts to develop modelling approaches that are efficient to predict the odor quality of complex odorants' mixtures such as those elicited by food products. The literature review suggests that molecular structure and/or olfactometry data would be an interesting basis for the modelling approach and that expert knowledge integration may provide an innovative strategy to cope with complex perceptual interactions. The issue is to develop a modelling strategy that can support the integration of these heterogeneous data. Finally, it is expected that the developed models can support a real breakthrough in the research field of flavor analysis and perception.

THESIS AIM AND OVERALL STRATEGY

The challenge of the present PhD project is to contribute to increase the efficiency of the food flavor analysis procedure by taking into account knowledge on odor mixture perception along the analytical path. The central aim is thus to develop a grey-box modelling approach able to predict the odor quality of complex mixtures of odorants into semantic descriptors such as sensory profiles of food odors.

Throughout the literature review, various food products were mentioned but the focus was set on wines. Indeed, there are numerous studies on wines' composition and their related odor. Hence, the odor prediction of such complex matrices is of interest. Thus, the PhD work was applied to 16 red wines. The choice of focusing on wines was motivated by the willingness to study a complex real food product and the availability of the chemical and sensory data on the 16 wines in the laboratory.

In the following Chapter (Chapter 2), the data collection and methodologies used are presented. First, the characterization of the 16 wines used along the thesis work is described in a data paper (Villière et al., **Data paper included in this manuscript**). Second, the method to collect flavorists' expertise is described. Then, two modelling methods relying respectively on an ontology and a fuzzy logic formalization are detailed.

The PhD work was divided into two distinct but complementary axes.

First axis: Molecular structure approach

The literature review highlights several predictive approaches based on structural parameters of odorants. One study has been applied to odorants' mixtures to predict perceptual similarity between mixtures (Snitz et al., 2013). In Chapter 3, an article is presented (Roche et al., **Article 1 of the manuscript**) to describe the application and improvement of the Snitz's model using a benchmark dataset containing elemental and configural mixtures of different level of complexity. The upgraded model was then applied to real food mixtures, namely to the 16 wines for which the predicted similarity was compared to the one inferred from experimental sensory profile data.

Second axis: Expertise integration approach

Chapter 4 focuses on the collection and representation of flavorists' expertise. The article 2 (Roche et al., **Article 2 of this manuscript**) presents the development of the Ontology for the Odor Perceptual Space (OOPS) that aimed to fix the vocabulary and identify the properties and relations between three lexical concepts: odor descriptors used for odorants and available in several databases, odor qualities used by experts and odor sensory attributes used in sensory profiles. The ontology approach was selected because it is suitable to formalize explicit experts' knowledge on the relationships within the odor perceptual space. OOPS was implemented in a computer interface and was tested as a predictive tool to estimate wines odor profiles.

In Chapter 5, an integrated modelling strategy, which combines chemical analysis results obtained on the 16 wines with expert flavorists knowledge as formalized in the OOPS ontology is described in a dedicated article (Roche et al., **Article 3 of this manuscript**). The modelling approach relies on fuzzy logic and optimization and was coded in MATLAB. Fuzzy logic is a methodology that allows to build a set of fuzzy rules of odorants associations to produce a given odor percept. The rules were optimized using a suitable genetic algorithm. This strategy was applied to predict the odor profiles of the 16 wines, which were compared to experimental sensory profile data.

CHAPTER 2: DATA COLLECTION & METHODOLOGIES

I. Introduction

In this chapter we firstly describe how we collected and organized the data used in the modelling approaches. Two types of data were collected and are detailed: food odor data characterizing 16 red wines from two grape varieties and expertise from flavorists. Then, two modelling methods relying respectively on an ontology and a fuzzy logic formalization are explained.

II. Food odor data

The study of wines' odor along with their aroma and taste is of importance because the consumption of wine is driven by its flavor (Lee & Noble, 2003).

Two approaches are suitable for wine odor analysis: chemical analysis and sensory analysis (Le Fur et al., 2003). As part of the project INNOVAROMA, both analyses were performed on 16 red-wines. The methods used to characterize the flavor of these wines were submitted as a **DATA PAPER** and are presented in the following pages.

The results obtained were compiled in a single dataset and made available on an open-access repository (Villière et al., 2018). The data further used in the manuscript were extracted from this dataset.

1. DATA PAPER

Characterization of the flavor of two red wine varieties using sensory descriptive analysis, volatile organic compounds quantitative analysis by GC-MS and odorant composition by GC-MS-O

Angélique Villière[1]‡, Ronan Symoneaux[2]‡, **Alice Roche**[3]‡, Aïda Eslami[4,5], Nathalie Perrot[6], Yves Le Fur[3], Carole Prost[1], Philippe Courcoux[4], Evelyne Vigneau[4], Thierry Thomas-Danguin[3]*, Laurence Guérin[7]*

1. GEPEA, Flavor group, Oniris, Univ. Bretagne Loire, CNRS, Nantes, France
2. USC 1422 GRAPPE, INRA, Ecole Supérieure d'Agricultures, Université Bretagne Loire, SFR 4207 QUASAV, Angers, France
3. CSGA, UMR1324 INRA, UMR6265 CNRS, AgroSup Dijon, Université de Bourgogne Franche - Comté, Dijon, France
4. StatSC, Oniris, Univ Bretagne Loire, INRA, Nantes, France
5. Centre for Heart Lung Innovation, University of British Columbia, Vancouver, British Columbia, Canada
6. GMPA, UMR782 INRA, AgroParisTech, Université Paris Saclay, Thiverval-Grignon, France
7. Institut Français de la Vigne et du Vin, Amboise, France

‡ co-first authors

* Correspondence to be sent to: Thierry Thomas-Danguin thierry.thomas-danguin@inra.fr or Laurence Guérin laurence.guerin@vignevin.com

This manuscript is currently submitted to Data in Brief.

Keywords: Wine, Descriptive sensory analysis, VOC, Olfactometry, GC-MS-O

Abstract

This paper describes data that were collected on 2 sets of 8 French red wines from two grape varieties, Pinot Noir (PN) and Cabernet Franc (CF). It provides, for the 16 wines, (i) sensory descriptive data obtained with a trained panel, (ii) volatile organic compounds (VOC) quantification data obtained by gas chromatography - mass spectrometry (GC-MS) and (iii) odorant composition obtained by gas chromatography - mass spectrometry - olfactometry (GC-MS-O). The raw data are hosted on an open-access research data repository (Villière et al., 2018).

Specifications Table

Subject area	Food science
More specific subject area	Wine research
Type of data	Microsoft Excel Worksheet containing 8 sheets (Information, Experimental factors, List sensory descriptors, Sensory descriptive analysis, List VOC, VOC quantification and GC-MS-O)
How data were acquired	• Sensory descriptive analysis The intensity of 33 sensory descriptors was rated by 16 trained panelists • VOC quantification Volatile compounds in wines were analyzed using gas chromatography coupled with mass spectrometer (GC-MS) • Odorant composition Odor-active compounds were identified using gas chromatography coupled with mass spectrometry and olfactometry (GC-MS-O)
Data format	Table in raw format (.xlsx)
Experimental factors	The experimental factors are: the grape variety, the vintage and the Protected Designation of Origin (PDO) of the wines
Experimental features	• Sensory descriptive analysis Sensory odor profile of the wines • VOC quantification Quantified the volatile compounds from the GC chromatogram • GC-MS-O Quantified the volatile compounds and their odor from the olfactometry and identified the compounds from the GC chromatogram
Data source location	France
Data accessibility	The raw data, provided as a Microsoft Excel Worksheet, are available on the Zenodo open-access research data repository http://doi.org/10.5281/zenodo.1213610

Value of the Data

- The data can help researchers to link sensory qualities of wines to their chemical composition (Vigneau et al., 2015).
- The data can be used as a benchmark to develop methods and tools to predict the odor of wines (Roche et al., 2017).
- The data can be compared to other wines varying in grape variety and vintage.

Data

The dataset gathers, for the 16 wines from two grape varieties, 4 blocks of data: (1) the experimental factors (the grape variety, the vintage and the protected designation of origin; Table 1), (2) the sensory descriptive data obtained with a trained panel using 33 sensory descriptors (Table 2), (3) the volatile organic compounds (VOC) quantification data obtained for 45 target odorants by gas chromatography - mass spectrometry (Table 3) and (4) the composition data, in terms of odor-active compounds, obtained by gas chromatography - mass spectrometry - olfactometry (Table 4).

Experimental Design, Materials, and Methods

Wines

Two sets of French red wines from two grape varieties, 8 Pinot Noir wines (PN) and 8 Cabernet Franc wines (CF) were analyzed (Table 1). The wines were selected out of 40 wines previously studied (Loison et al., 2015). The main factors allowed for were vintage (2009 and 2010) and protected designation of origin (PDO).

Table 1: Wines experimental factors.

Wine	Grape_variety	Vintage	PDO
PN1	Pinot Noir	2010	Bourgogne
PN2	Pinot Noir	2009	Bourgogne
PN3	Pinot Noir	2009	Bourgogne
PN4	Pinot Noir	2009	Bourgogne Hautes Côtes de Beaune
PN5	Pinot Noir	2009	Savigny-lès-Beaune
PN6	Pinot Noir	2010	Maranges
PN7	Pinot Noir	2009	Côte de Nuits-Villages
PN8	Pinot Noir	2009	Ladoix
CF1	Cabernet Franc	2010	Bourgueil
CF2	Cabernet Franc	2010	Chinon
CF3	Cabernet Franc	2009	Chinon
CF4	Cabernet Franc	2010	St-Nicolas-de-Bourgueil
CF5	Cabernet Franc	2010	Bourgueil
CF6	Cabernet Franc	2010	Bourgueil
CF7	Cabernet Franc	2010	Bourgueil
CF8	Cabernet Franc	2010	Saumur

Sensory descriptive analysis

The sensory descriptive analysis of the 16 wines was performed at Groupe ESA, USC GRAPPE Senso'Veg (Angers, France).

Wines preparation

The wines were opened 30 minutes before the sensory evaluation and served (5 cL) in white ISO wine tasting glasses (ISO 3591:1977 / Sensory analysis - Apparatus - Wine-tasting glass) at room temperature.

Sensory evaluation

Sixteen trained panelists, 6 women and 10 men (age range 35-71), participated in the study. Before the sensory descriptive experiment, the judges were trained in 17 training sessions of 1 hour each. This training consists in familiarization with the task and the vocabulary and a selection of sensory descriptors applied to the wines. During the familiarization step, the panelists did odor recognition tests on testing strip and on wines to become familiar with the specific vocabulary of the sensory descriptors of wines and smelled different standard aromatic references. During the sensory descriptors selection, the panelists were provided

with an initial list of 84 descriptors. The list was elaborated by compiling terms from other lists employed in the description of wines from different varieties and geographical origins. Descriptors were arranged in the list by odor families: animal, burnt, floral, fruity, herbaceous, mineral, nut, spicy, undergrowth and others. Panelists modified the initial list of terms by eliminating those terms they considered irrelevant, ambiguous, or redundant and by adding attributes they considered pertinent while describing 15 wines of similar characteristics (grape variety and origin) as those of the study. Furthermore, those terms cited by less than 15% of the panel were eliminated from the list. At the end of the training, the list included 33 descriptors (Table 2).

Table 2: Sensory descriptors used by the trained panel for the sensory descriptive analysis.

Artichoke	Cherry fresh	Leather	Strawberry fresh
Bell pepper	Cherry stone	Musk	Toasty
Blackberry fresh	Clove	Pepper	Undergrowth
Blackcurrant bud	Cut grass	Plum cooked	Vanilla
Blackcurrant fresh	Elderflower	Plum fresh	Violet
Blueberry fresh	Ethanol	Prune	Woody
Brioche	Firestone	Raspberry fresh	
Butter	Geranium	Smoky	
Cherry cooked	Hay	Strawberry cooked	

During the sensory descriptive experiment, the judges had to evaluate monadically the 16 wines (orthonasal and retronasal olfaction) and to rate the intensity of 33 sensory descriptors on 14 cm linear scales; ratings were transformed into scores from 0 to 10. The protocol consisted in 3 repetitions by panelist for the orthonasal olfaction and 2 repetitions by panelist for the retronasal olfaction and a randomization of the presentation order of the wines according to the Williams Latin square.

The configuration of the set of wines evaluated through orthonasal olfaction is represented after Principal Components Analysis (PCA) in Figure 1.

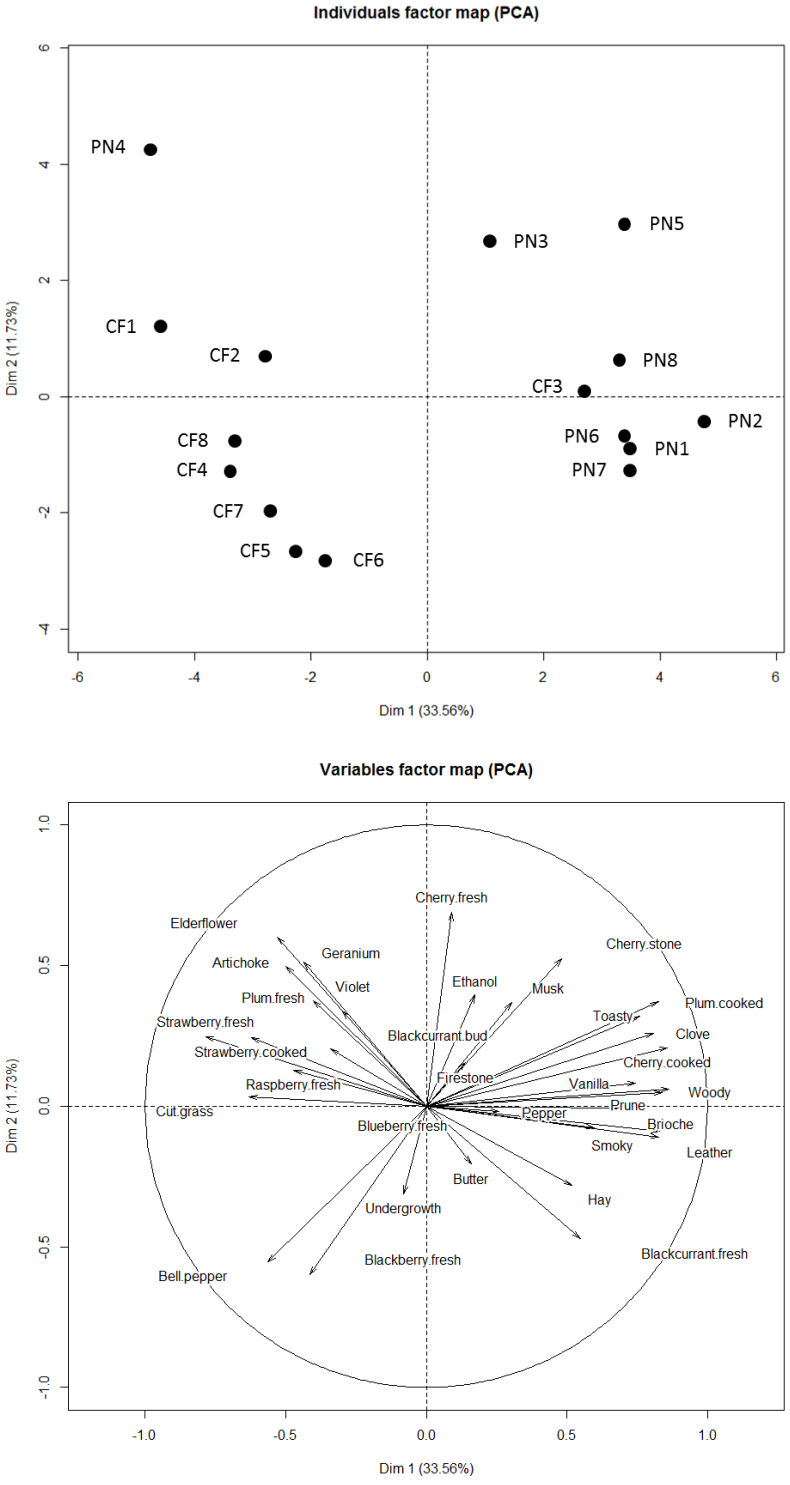


Figure 1: PCA maps based on the two first dimensions illustrating the configuration of the 16 wines (individuals) evaluated orthonasaly during sensory profiling on 33 sensory descriptors (variables). Mean for each sensory descriptor was taken into account.

Volatile organic compounds quantitative analysis

The 16 wines were analyzed by GC-MS to quantify 45 target compounds. The odorants listed in Table 3 were analyzed by a subcontracting external laboratory.

Table 3: Volatile organic compounds (VOC) quantified by GC-MS analysis and their corresponding CAS number.

VOC	CAS number	VOC	CAS number
1-Hexanol	111-27-3	Ethyl acetate	141-78-6
1-Octanol	111-87-5	Ethyl butyrate	105-54-4
1-Phenoxy-2-propanol	770-35-4	Ethyl caproate	123-66-0
2,3-Butanedione	431-03-8	Ethyl isobutyrate	97-62-1
2-Ethylhexan-1-ol	104-76-7	Ethyl isovalerate	108-64-5
2-Isobutyl-3-methoxypyrazine	24683-00-9	Ethyl lactate	97-64-3
2-Methyl-1-butanol	137-32-6	Ethyl octanoate	106-32-1
2-Methylbutyl acetate	624-41-9	Ethyl propionate	105-37-3
2-Phenylethanol	60-12-8	Furaneol	3658-77-3
3-Methyl-1-butanol	123-51-3	Hexyl acetate	142-92-7
4-Ethyl-2-methoxyphenol	2785-89-9	Homofuraneol	27538-10-9
4-Ethylphenol	123-07-9	Isoamyl acetate	123-92-2
Acetaldehyde	75-07-0	Isoamyl propionate	105-68-0
Acetic acid	64-19-7	Isovaleric acid	503-74-2
alpha-Ionone	127-41-3	Methional	3268-49-3
beta-Ionone	79-77-6	Methionol	505-10-2
Butyl acetate	123-86-4	Pentyl propionate	624-54-4
Butyric acid	107-92-6	Phenol	108-95-2
Damascenone	23726-93-4	Phenylacetaldehyde	122-78-1
Dimethyl Sulfide	75-18-3	Phenylacetic acid	103-82-2
Ethyl 2-methylbutyrate	7452-79-1	Propionic acid	79-09-4
Ethyl 3-hydroxybutyrate	5405-41-4	trans-3-Hexen-1-ol	544-12-7
Ethyl 6-hydroxyhexanoate	5299-60-5		

After adding a standard to the wines, they were acidified and salt saturated. The extraction of the volatile compounds was realized with a SPME fiber (CAR-DVB-PDMS) placed in the headspace of the vial for 60 min at 45°C. Extracts were analyzed by GC (Shimadzu 2010) coupled with a mass spectrometer (Shimadzu QP2010+). Volatile compounds were desorbed in the injection port (splitless/split) of the GC and separated on a PEG modified column (DB-FFAP 30m × 0.32mm × 0.25 µm). Mass spectra were recorded in electron impact mode (70

eV) with a scan/SIM scanning method. The identification of Acetaldehyde, Dimethyl sulfide, Ethyl acetate, Acetic acid, 2-Ethylhexan-1-ol, Propionic acid and Phenol were performed by comparison with mass spectra from the literature (WILEY257, NIST, AROMALYSIS databases). Their quantification is based on an internal calibration by isotopic dilution with the compounds ethanal-13C2, dimethylsulfide-d6, ethyl acetate-13C2, acetic acid-d4, 2-Ethylhexan-1-ol-d17, propionic acid-d5 and phenol-d6. The identification and quantification of all other compounds are based on a calibration of the method with these same reference compounds.

Analysis of wines by GC-MS-O

The 16 wines were analyzed by GC-MS-O at ONIRIS, UMR CNRS 6144 GEPEA Flavor group (Nantes, France).

Extraction methods

The wines were firstly oxygenated by a Venturi aerator, and then 7 mL of wine was poured in a 22 mL vial tightly capped with a Teflon/silicon septum. Volatile compounds from the wine samples were extracted by a representative procedure (Villière et al., 2012). Prior to extraction, vials were incubated at 34°C for 1 hour. After that, volatile compounds were extracted by headspace solid phase micro-extraction with a Car/PDMS fiber (10 mm length, 85 µm film thickness; Supelco, Bellefonte, PA, USA) placed in the headspace of the vial for 10 min at 34°C.

Chromatographic conditions

The extracts were analyzed by GC (Agilent Technologies 6890N, Wilmington, DE, USA) coupled with a quadripolar mass spectrometer (Agilent Technologies, 5973 Net- work), FID and sniffing port (ODP2, Gerstel, Baltimore, MD, USA) to identify odorant compounds.

Volatile compounds were desorbed in the injection port of the GC (T: 260°C; splitless mode for 5 min) and separated on a DB-Wax column (length: 30 m, internal diameter: 0.25 mm, film thickness: 0.5 µm). Hydrogen was used as carrier gas at constant flow (1 $mL.min^{-1}$). The oven temperature program was set from 50°C (0 min) to 80°C at 5°C min^{-1}, from 80°C to 200°C at 10°C min^{-1} and from 200 to 240°C (4 min) at 20°C min^{-1}. Effluent from the end of the GC column was split 1:1:1 between the MS, the FID (250°C, air/H2 flow: 450/40 $mL.min^{-}$

[1]), and the sniffing port. Peaks were integrated with MSD Chemstation software (Agilent Technologies). Mass spectra were recorded in electron impact mode (70 eV) between 33 and 300 m/z mass range at a scan rate of 2.7 scan.s^{-1}.

Olfactometry

GC effluent was carried to the sniffing port using a deactivated and uncoated fused silica capillary column, heated to 200°C. The sniffing port was supplied with humidified air at 40°C with a flow of 600 mL.min^{-1}.

Olfactometry was conducted by eight judges experienced in GC-O analysis. They were asked to express their perceptions via the olfactometric software interface, representing an aroma wheel designed for wine analysis with 56 descriptors (Villière et al., 2015). Characteristics of the perceptions were recorded throughout each judge's analysis and results were directly obtained from the olfactometric software.

Odorant compounds identification

The identification of compounds corresponding to each odorant zone was performed by comparing linear retention index and mass spectra of detected compounds with those of the databases (Wiley 6.0 and in-house databases), by injection of the standard compounds when available, and by comparison of the odor perceived with those referenced in databases (in house database and The good scents company database (Luebke, 1980)). The results of the identification are shown in Table 4 with the list of 49 odorants. Compounds non-identified were named after their apex indice number (34 non-identified compounds) (Villière et al., 2018).

Table 4: Odorant compounds identified by GC-MS-O analysis and their corresponding CAS number.

Odorant	CAS number	Odorant	CAS number
1-Octen-3-one	4312-99-6	Dimethyl sulfide	75-18-3
2,3-Butanedione	431-3-8	Ethanol	64-17-5
2,3-Pentanedione	600-14-6	Ethyl acetate	141-78-6
2,6-Dimethoxyphenol	91-10-1	Ethyl butanoate	105-54-4
2-Methoxyphenol	90-05-1	Ethyl decanoate	110-38-3
2-Methylpropyl acetate	110-19-0	Ethyl dodecanoate	106-33-2
3-Ethylphenol	620-17-7	Ethyl hexanoate	123-66-0
3-Isobutyl-2-methoxypyrazine	24683-00-9	Ethyl octanoate	106-32-1
3-Isopropyl-2-methoxypyrazine	25773-40-4	Ethyl propanoate	105-37-3
3-Mercapto-1-hexanol	51755-83-0	Ethyl-2-methylbutanoate	7452-79-1
3-Methyl-1-butanol	123-51-3	Ethyl-2-methylpropanoate	97-62-1
3-Methylbutanal	590-86-3	Ethyl-3-methylbutanoate	108-64-5
3-Methylbutyl acetate	123-92-2	Hexanoic acid	142-62-1
4-Ethyl guaïacol	2785-89-9	Isovaleric acid	503-74-2
4-Ethylphenol	123-07-9	m-Cresol	108-39-4
4-Methyl-1-pentanol	626-89-1	Methanethiol	74-93-1
Acetaldehyde	75-07-0	Methional	3268-49-3
Acetic acid	64-19-7	Methionol	505-10-2
Benzaldehyde	100-52-7	Methyl-2-methylpropenoate	80-62-6
Benzene acetaldehyde	122-78-1	p-Cresol	106-44-5
Benzene ethanol	60-12-8	Phenethyl acetate	103-45-7
Benzene methanol	100-51-6	Phenol	108-95-2
Butyric acid	107-92-6	Sulphur dioxide	7446-09-5
Butyrolactone	96-48-0	Whyskeylactone	39212-23-2
Decanoic acid	334-48-5		

Acknowledgments

The authors would like to thank the panelists involved in the sensory evaluation and the olfactory task. They also thank Aurore Boissy, Erell Bourhis, Caroline Chiffoleau, Solenn Jourdren, Catherine Fillonneau, Laurent Lethuaut, InterLoire (Interprofession des vins du Val de Loire) and the BIVC (Bureau Interprofessionnel des Vins du Centre-Loire) for their assistance at different stages of the project.

Funding sources

These data were collected as part of the INNOVAROMA research program on wine, which was conducted with financial support from the regional councils of Pays de la Loire, Centre, Bretagne, and Bourgogne (DINOS-AGRALE-12-2011, DINOS-AGRALE-12-2012, PARI-DINOS-2013).

References

Loison, A, Symoneaux, R, Deneulin, P, Thomas-Danguin, T, Fant, C, Guérin, L & Le Fur, Y. 2015. Exemplarity measurement and estimation of the level of interjudge agreement for two categories of French red wines. Food quality and preference, 40:240–251. doi:10.1016/j.foodqual.2014.10.001.

Luebke, W. 1980. The good scents company. Available at: http://www.thegoodscentscompany.com/index.html

Roche, A, Perrot, N, Chabin, T, Villière, A, Symoneaux, R & Thomas-Danguin, T. 2017. In silico modelling to predict the odor profile of food from its molecular composition using experts' knowledge, fuzzy logic and optimization: Application on wines. 2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Montréal, Canada, 28-31 May. doi:10.1109/ISOEN.2017.7968875.

Vigneau, E, Courcoux, P, Lefebvre, R, Villière, A & Symoneaux, R. 2015. Regression trees and random forests as a tool for identifying the volatile organic compounds implied in the olfactory perception of wines. 11th Pangborn Sensory Science Symposium, Göteborg, Sweden, 23-27 August.

Villière, A, Arvisenet, G, Lethuaut, L, Prost, C & Sérot, T. 2012. Selection of a representative extraction method for the analysis of odourant volatile composition of French cider by GC-MS-O and GC × GC-TOF-MS. Food Chemistry, 131:1561-1568. doi:10.1016/j.foodchem.2011.10.008.

Villière, A, Le Roy, S, Fillonneau, C, Guillet, F, Falquerho, H, Boussely, S & Prost, C. 2015. Evaluation of aroma profile differences between sué, sautéed, and pan-fried onions using an innovative olfactometric approach. Flavour, 4(24):1-19. doi:10.1186/s13411-015-0034-0.

Villière, A, Symoneaux, R, Roche, A, Eslami, A, Perrot, N, Le Fur, Y, Prost, C, Courcoux, P, Vigneau, E, Thomas-Danguin, T & Guérin, L. 2018. Dataset on the characterization of the flavor of two red wine varieties using sensory descriptive analysis, volatile organic compounds quantitative analysis by GC-MS and odorant composition by GC-MS-O. Zenodo [dataset], doi:10.5281/zenodo.1213610.

2. Summary of the results

To sum up, we obtained chemical analysis and sensory analysis data from the characterization of eight red wines from Pinot Noir and eight red wines from Cabernet Franc.

From the chemical analysis data we obtained the list of odorants contained in each wine (total of 49 odorants). Moreover the data allows the calculation of the relative intensity of each odorant by calculating their nasal impact frequency (NIF) score. NIF are calculated as the ratio of panelists who smelled the odorants on the total number of panelists (total = 8).

From the sensory data, we focused on the orthonasal evaluation of the wines, so we have 3 repetitions for each panelist and for each wine. We assessed the panel's performance through its capabilities to use efficient descriptors to discriminate between wines odor profile. In Table 2.1, the odor descriptors are ranked according to p-values associated with the F-test of ANOVA. Among the 33 sensory descriptors studied, 8 were identified as discriminant between the wines. On the one hand, results highlighted also the good repeatability of the panelists for the three repetitions, except for the descriptor Plum Fresh ($p = 0.03$). The low intra-individual variability showed that the training of the panelists was efficient. On the other hand, inter-individual variability was shown with a significant panelist effect for all the descriptors. In the following part of the work, we focused on the 17 more discriminant odor descriptors ($p < 0.25$) from now called odor sensory attributes (OSA). The OSA Ethanol was not taken into account because it corresponds to the name of a molecule rather than an odor percept such as Rummy or Wine-like.

Table 2.1: Discrimination of the wines according to 33 sensory descriptors and panelist's performance assessed by the p-values associated with the F-test for each descriptor. For the 33 descriptors, the column "Wine" reports the p-value associated to the factor Wine in ANOVA and reflects the discrimination ability of the descriptors for the wines. The column "Panelists" reports the p-value associated to the factor Panelist in ANOVA and reflects the inter-individual variability for each descriptor. The column "Repetition" reports the p-value associated to the factor Repetition in ANOVA and reflects the intra-individual variability within the three repetitions. Significance is shown with the orange shading ($p<0.05$). Sensory descriptors in bold correspond to the odor sensory attributes (OSA) kept further in modelling studies.

	Wine	Panelist	Repetition
Woody	2.00E-08	5.77E-80	0.25
Smoky	2.25E-04	2.00E-24	0.59
Cherry cooked	8.71E-04	2.97E-86	0.85
Vanilla	1.45E-03	2.09E-131	0.97
Strawberry fresh	4.60E-03	1.06E-14	0.4
Musk	1.32E-02	1.35E-28	0.07
Violet	3.00E-02	2.70E-12	0.21
Leather	3.97E-02	5.77E-17	0.57
Cherry stone	5.16E-02	2.09E-37	0.42
Blackcurrant bud	5.37E-02	1.71E-45	0.61
Toasty	5.66E-02	4.50E-24	0.28
Artichoke	6.11E-02	5.02E-22	0.68
Prune	7.94E-02	6.57E-27	0.38
Cut grass	8.37E-02	3.22E-17	0.05
Ethanol	1.08E-01	7.60E-81	0.75
Blackcurrant fresh	1.14E-01	1.18E-32	0.32
Bell pepper	1.64E-01	4.00E-58	0.17
Cherry fresh	2.16E-01	1.73E-37	0.61
Raspberry fresh	3.27E-01	2.87E-08	0.06
Brioche	3.43E-01	3.55E-21	0.77
Hay	3.59E-01	1.01E-52	0.14
Clove	4.15E-01	7.54E-10	0.61
Plum cooked	4.48E-01	6.98E-31	0.99
Elderflower	4.55E-01	8.21E-16	0.83
Blackberry fresh	4.68E-01	1.26E-29	0.53
Geranium	5.51E-01	3.51E-10	0.61
Undergrowth	6.51E-01	2.52E-06	0.34
PlumFresh	6.62E-01	1.31E-09	0.03
Strawberry cooked	6.82E-01	6.21E-47	0.8
Butter	6.95E-01	3.65E-24	0.75
Firestone	8.49E-01	2.87E-21	0.93
Blueberry fresh	8.55E-01	1.83E-21	0.25
Pepper	8.89E-01	5.28E-51	0.74

III. Expertise

One of the innovative aspects of this work was to integrate flavor expertise knowledge into a predictive modelling strategy. Two types of expertise were collected: construction of the mixture rules for odor sensory attributes (OSA) and relationships between odor descriptors (OD) used for odorants' description and odor qualities (OQ) used by expert flavorists. Thus, we gathered knowledge from three levels of the odor perceptual space: OD, OQ and OSA.

1. Odor sensory attributes (OSA)

The expertise of four senior flavorists was collected. Each expert was invited to participate in a semi-directed interview by phone for two sessions of 1 hour following the elicitation process described below. In the first session, experts were not aware of the studied food matrix (i.e. wine) and thus we collected generic knowledge. In the second session, experts were informed of the studied food matrix (data not used).

a. Elicitation process

Experts were presented the list of the 17 OSA selected from the sensory profiles of the 16 red wines previously characterized. They were asked, for each OSA, to tell if the OSA was composed of a single OQ or a combination of OQ. In the latter case, experts were asked to list the OQ they may use to construct the perception of the OSA and to precise the proportion of each OQ; they freely used their own OQ. Proportions were collected on a symbolic scale +++ (very high), ++ (high), + (medium), +/- (weak), near 0 (trace) further translated in a numeric scale (+++ = 3, ++ = 2, + = 1, +/- = 0.5, near 0 = 0.25). The numerical data were then converted into row-wise percentages to obtain proportions of OQ for each expert (for example, the value 0.33 obtained for the OQ Almond for the expert E1 is the result of $\frac{1}{1+2}$, 1 being the numerical value of the OQ Almond and 1+2 the sum of the values from the two OQ used by the expert).

b. Results

As for example, the data collected for the OSA Prune are presented in Table 2.2. The symbolic data, numerical data and proportions are detailed. The results highlighted several

combinations of OQ able to elicit the perception of the OSA Prune. Indeed, the OSA Prune was described as the combination of the OQ "Almond and Lactonic" for the expert E1, "Cooked, Fruity, Honey, and Lactonic" for the expert E2, or "Cooked and Fruity" for the expert E4. In this case, the expert E3 did not answer the question.

Table 2.2: Combination of odor qualities (OQ) eliciting the perception of the odor sensory attribute (OSA) Prune. E1 to E4 refers to the four flavorists interviewed.

Symbolic data					
Experts	Almond	Cooked	Fruity	Honey	Lactonic
E1	+	*Not used*	*Not used*	*Not used*	++
E2	*Not used*	++	+	Near 0	+
E3	*Missing data*	*Missing data*	*Missing data*	*Missing data*	*Missing data*
E4	*Not used*	+++	+	*Not used*	*Not used*
Numerical data					
Experts	Almond	Cooked	Fruity	Honey	Lactonic
E1	1	*Not used*	*Not used*	*Not used*	2
E2	*Not used*	2	1	0.25	1
E3	*Missing data*	*Missing data*	*Missing data*	*Missing data*	*Missing data*
E4	*Not used*	3	1	*Not used*	*Not used*
Proportions					
Experts	Almond	Cooked	Fruity	Honey	Lactonic
E1	0.33	*Not used*	*Not used*	*Not used*	0.67
E2	*Not used*	0.47	0.24	0.06	0.24
E3	*Missing data*	*Missing data*	*Missing data*	*Missing data*	*Missing data*
E4	*Not used*	0.75	0.25	*Not used*	*Not used*

The proportions of OQ corresponding to the data collected from the expert flavorists for all the 17 OSA are presented in Table 2.3.

Table 2.3: Combination of odor qualities (OQ) eliciting the perception of 17 odor sensory attributes (OSA). E1 to E4 refers to the four flavorists interviewed. Nu: Not used. Md: Missing data.

Artichoke							
E1	Md						
E2	Md						
E3	Md						
E4	Md						

Cut-grass							
	Cut-grass						
E1	**1**						
E2	**1**						
E3	**1**						
E4	**1**						

Leather							
	Leather	Animal	Coumarinic	Smoky			
E1	**1**	*Nu*	*Nu*	*Nu*			
E2		0.5	0.5	*Nu*			
E3	**1**	*Nu*	*Nu*	*Nu*			
E4		0.33	*Nu*	0.67			

Musk							
E1	Md						
E2	Md						
E3	Md						
E4	Md						

Smoky							
	Smoky						
E1	*Md*						
E2	**1**						
E3	**1**						
E4	**1**						

Toasty							
	Toasty	Cooked					
E1	**1**						
E2	0.67	0.33					
E3	**1**						
E4	**1**						

Vanilla							
	Vanilla						
E1		10 OQ (Almond, Bean, Buttery, Caramelic, Coumarinic, Creamy, Floral, Rummy, Spicy, Vanilla)					
E2	**1**						
E3	*Md*						
E4	**1**						

Violet							
	Violet						
E1	**1**						
E2	**1**						
E3	**1**						
E4	**1**						

Woody							
	Woody						
E1	Md						
E2	**1**						
E3	**1**						
E4	**1**						

Bell pepper							
	Floral	Fruity	Green	Sulfurous	Toasty	Vegetable	
E1	*Nu*	*Nu*	0.8	0.2	*Nu*	*Nu*	
E2	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*	
E3	0.17	0.33	*Nu*	0.17	*Nu*	0.33	
E4	*Nu*	*Nu*	0.25	0.25	0.5	*Nu*	

Blackcurrant bud							
	Floral	Fresh	Fruity	Green	Sulfurous	Vanilla	Wine-like
E1	0.22	*Nu*	*Nu*	*Nu*	0.11	*Nu*	*Nu*
E2	0.1	*Nu*	*Nu*	0.2	0.4	*Nu*	*Nu*
E3	*Nu*	0.25	0.25	*Nu*	0.12	*Nu*	0.38
E4	*Nu*	0.33	0.33	*Nu*	0.33	*Nu*	*Nu*

Blackcurrant fresh							
Experts	Floral	Fresh	Fruity	Green	Sulfurous	Wine-like	
E1	0.13	0.53	0.27	*Nu*	0.07	*Nu*	
E2	0.11	*Nu*	0.44	0.22	0.22	*Nu*	
E3	*Nu*	0.38	0.25	*Nu*	0.13	0.25	
E4	*Nu*	0.5	0.25	*Nu*	0.25	*Nu*	

Cherry cooked							
	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
E1	0.24	0.28	0.07	*Nu*	0.21	*Nu*	0.21
E2	0.29	0.14	*Nu*	0.57	*Nu*	*Nu*	*Nu*
E3	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*
E4	0.13	0.38	*Nu*	0.38	0.06	0.06	*Nu*

Cherry fresh							
	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
E1	0.26	0.22	0.07	*Nu*	0.22	*Nu*	0.22
E2	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*
E3	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*
E4	0.12	0.24	*Nu*	0.35	0.24	0.06	*Nu*

Cherry stone							
	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
E1	0.39	0.18	0.06	*Nu*	0.18	*Nu*	0.18
E2	1	*Nu*	*Nu*	*Nu*	*Nu*	*Nu*	*Nu*
E3	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*	*Md*
E4	0.25	0.25	*Nu*	0.38	0.06	0.06	*Nu*

Prune							
	Almond	Cooked	Fruity	Honey	Lactony		
E1	0.33	*Nu*	*Nu*	*Nu*	0.67		
E2	*Nu*	0.47	0.24	0.06	0.24		
E3	*Md*	*Md*	*Md*	*Md*	*Md*		
E4	*Nu*	0.75	0.25	*Nu*	*Nu*		

Strawberry fresh							
	Cooked	Floral	Fruity	Green			
E1	*Md*	*Md*	*Md*	*Md*			
E2	0.17	0.33	0.5	*Nu*			
E3	*Md*	*Md*	*Md*	*Md*			
E4	0.4	*Nu*	0.33	0.27			

The OSA Cut-grass, Leather, Smoky, Toasty, Vanilla, Violet and Woody were considered as composed of 1 OQ for more than half of the experts and were further considered as so. We were not able to collect data concerning the OSA Artichoke, thus it was removed from the set of OSA to predict. The 8 remaining OSA were considered as composed of several OQ.

2. Link between odor descriptors (OD) and odor qualities (OQ)

As presented in the literature review, odorants can be described in verbal terms by a set of words, namely odor descriptors (OD), which can be found in several databases. The OD of the 49 odorants identified from the chemical analysis of the 16 red wines previously characterized were collected. We compiled the data of three databases: Arctander's handbook (3102 chemicals described by Steffen Arctander himself), Flavor-Base (commercially available Leffingwell & Associates database, marketed as Flavor-Base Pro © 2010, flavor descriptions collected from many sources over the course of more than 40 years), and The good scents company (publicly available database, the odor descriptions from one to several sources are listed in the "Organoleptic Properties" section). Then we aggregated the information of the three databases. For a given odorant, the set of OD was the union of the OD from the three databases. We ended up with a corpus of 175 different OD.

However, the odor dimension used by flavorists when they create a flavor is a set of odor qualities (OQ). Hence, we wanted to establish the relationships between these two dimensions of the olfactory space: OD and OQ. A junior flavorist from the ISIPCA (International institute for perfumery, cosmetics, and food flavors) was interviewed through a check-all-that-apply (CATA) questionnaire (Dooley et al., 2010). The CATA list consisted of the corpus of 20 OQ defined by the experts during the elicitation step. For the 175 OD mentioned in the paragraph above, the flavorist was asked if the OD supported none, one, or several OQ. For instance for the OD "Apple", the flavorist was asked to tick all the OQ that correspond to this OD (e.g. "Fruity").

These data collections (from databases and expertise) allowed us to translate the OD space into OQ space. The results were presented following an ontology approach.

IV. Ontology

In computer science, ontology is expressed as a formal representation of knowledge by a set of concepts within a domain and the relationship between these concepts (Gruber, 1993). Ontology has gained much importance not only in the field of artificial intelligence, but also in the fields of natural language processing (Medjkoune et al., 2016) or knowledge representation and acquisition (Ibanescu et al., 2016).

Ontology construction is an iterative process and involves two main steps. Firstly, the design and development step aims to specify the scope and purpose of the ontology and also reveals the relationship among classes and subclasses (taxonomic hierarchy). Secondly, the validation and feedback step are performed by experts. If modifications are identified by the experts, the changes are incorporated in the ontology.

1. Definitions

An ontology together with a set of classes and their properties constitutes a knowledge base, mathematically defined as a triplet {C, R, P} (Madalli et al., 2017), where C represents classes, R represents the hierarchical relations among classes and P represents the properties.

A class, also called concept (Noy & McGuiness, 2001), is defined as a group of individuals sharing some properties. Classes can be organized in a specialization hierarchy using sub-classes. Members of a subclass inherit the characteristics of their parent class. When two classes are disjoint, an individual cannot be a member of disjoint classes.

The hierarchical relations denote the relations between classes and between properties and are defined by relations such as is_a, subclass_of, value_of.

Properties are non-hierarchical relations and thus relate classes to other classes or individuals of one class to individuals of another class. Properties are divided in two groups: data property and relational property. For example, in the context of the domain food, color of wine or calorie content are data properties, whereas recipe of a meal is a relational property.

Furthermore, ontologies can be used to describe a set of data by assigning instances belonging to a given set I to the classes C (such as specific persons, places, things). In this case,

ontologies are called populated ontologies (James et al., 2010) and are defined by the set {C, R, P, I, g}, where I represents the instances and g is a function which associates a finite number of instances to each classes of C.

2. Example

A highly simplified example of an ontology for wines is graphically presented in Figure 2.1. It is a very limited extract of an ontology where the class "Wines" represents all wines. This class can have subclasses that represent concepts which are more specific than the superclass. Here, the class "Wines" is divided in the classes of red, rosé, and white wines. Specific wines are instances of these classes. Thus, Vosne-Romanée and Meursault wines are respectively instances of the class Red wines and White wines. Because Vosne-Romanée and Meursault wines are made in Burgundy and more specifically in Côte d'Or, both instances are related to the "Côte-d'Or" subclass of "Region". In this example, the knowledge was collected from the Bourgogne Wine Board's website (BIVB: *Bureau Interprofessionnel des Vins de Bourgogne*) and translated into an ontology formalism.

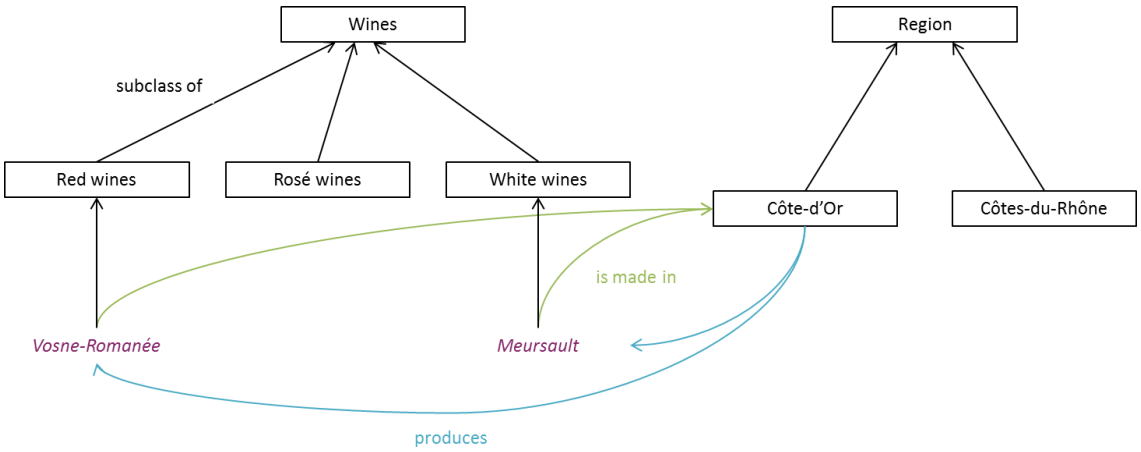


Figure 2.1: Graphical representation of very limited extract of an ontology in the wine domain. Black boxes represent classes, purple is used for instances. Arrows in black correspond to "subclass of" properties, green "is made in", and blue "produces". Figure adapted from Noy & McGuiness, 2001.

Ontologies define a common vocabulary for researchers who need to share information in a domain, for instance dairy gels (Ibanescu et al., 2016). Thus, they are valuable in sharing a common understanding of the structure of information among people or software agents.

V. Fuzzy logic

The term fuzzy logic was introduced with the 1965 proposal of fuzzy set theory by Zadeh. Unlike the conventional theory of sets, according to which an element either belongs (1) or does not belong (0) to a set, the approach of fuzzy logic imitates the way of decision making in humans that involves all intermediate possibilities within the interval [0,1] (Zadeh, 1965; Dubois & Prade, 1980).

The differences between classical logic and fuzzy logic are presented in Figure 2.2 through the exemple of the membership degrees between food liking and the linguistic term "delicious". Food liking correspond to a score on a 10-point scale. In classical logic, if a food product is given a score of 9 or higher, the food is part of the set "delicious", whereas food with lower score are not part of the "delicious" set. In fuzzy logic, food with a liking score starting at 7 are part of the set "delicious" with membership degrees varying from 0 to 1. A fuzzy set is then characterized by a function represented by a real number within the interval [0,1]. This membership function is useful to introduce graduality in different concepts: similarity, uncertainty, and constraint.

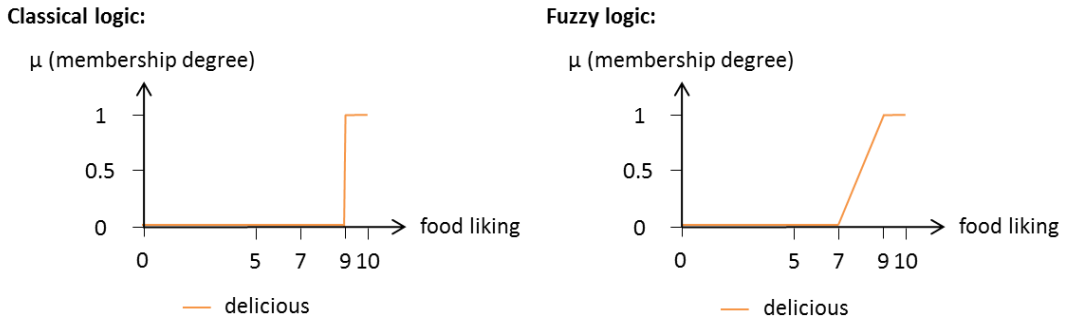


Figure 2.2: Comparison between classical and fuzzy logic.

The advantage of this formalism is to cope with human reasoning. Moreover, fuzzy functions can be created from expertise knowledge (Sicard et al., 2011).

1. Mathematical definitions of a fuzzy set

A fuzzy set E in universe of discourse U can be defined by Equation 2.1.

Equation 2.1:

$$E = \{(u, \mu_E(u)) \backslash u \in U\}$$

$$\mu_E : U \rightarrow [0,1]$$

μ_E is the membership function of the set E. It represents the set of membership grades $\mu_E(u)$ of a value u defined in a universe of discourse U, often numerical. The value of the membership grade is a real number within the interval [0,1], representing the membership degree of u to E.

The shape of the membership function is chosen arbitrarily by following the advice of the expert or by statistical studies: sigmoid, hyperbolic, tangent, exponential, Gaussian, or any other form can be used. Equation 2.2 and Figure 2.3 correspond to a trapezoidal representation.

Equation 2.2:

$$\mu_E(u) = \begin{cases} 0 \ (u \leq a_1) \\ \frac{u - a_1}{a_2 - a_1} \ (a_1 < u \leq a_2) \\ 1 \ (a_2 < u \leq a_3) \\ \frac{a_4 - u}{a_4 - a_3} \ (a_3 < u \leq a_4) \\ 0 \ (a_4 < u) \end{cases}$$

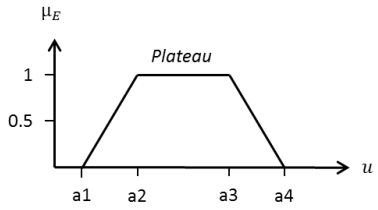


Figure 2.3: Fuzzy trapezoidal function. a2 and a3 represent the bounds of the plateau. a1 varies between 0 and a2, a4 varies between a3 and 1.

2. Fuzzy algorithm

Fuzzy logic systems are constructed with three main parts: a fuzzification module to transform the system inputs into fuzzy sets, fuzzy rules, and a defuzzification module to transform the fuzzy set obtained into output values. The parametrization of the fuzzy algorithm can be provided by experts (fuzzification and defuzzification, IF-THEN rules). The different parts of a fuzzy logic system are presented through a fuzzy system (adapted from Dernoncourt, 2011), which for instance can predict the tip given in a restaurant regarding the food liking and service quality. In this case, food liking and service quality are scored on a 10-point scale. The output of the fuzzy system is the tip given as the percentage of the bill (Figure 2.4).

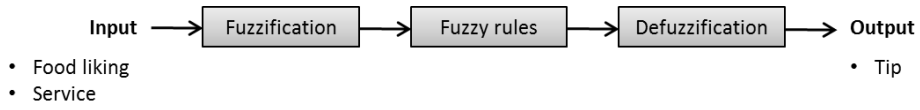


Figure 2.4: Fuzzy system architecture. Example of a fuzzy system for predicting the tip given in a restaurant regarding the food liking and service quality.

a. Fuzzyfication

The membership functions link fuzzy input and linguistic variables. Each linguistic variable is defined by a respective range and fuzzy set. When fuzzy sets are defined for a linguistic variable, the goal is not to define the linguistic variable exhaustively. On the contrary, the

focus is set only on the sets that will be useful later in the definition of the fuzzy rules. In Figure 2.5, the fuzzyfication of the variable food liking, service quality and tip are presented.

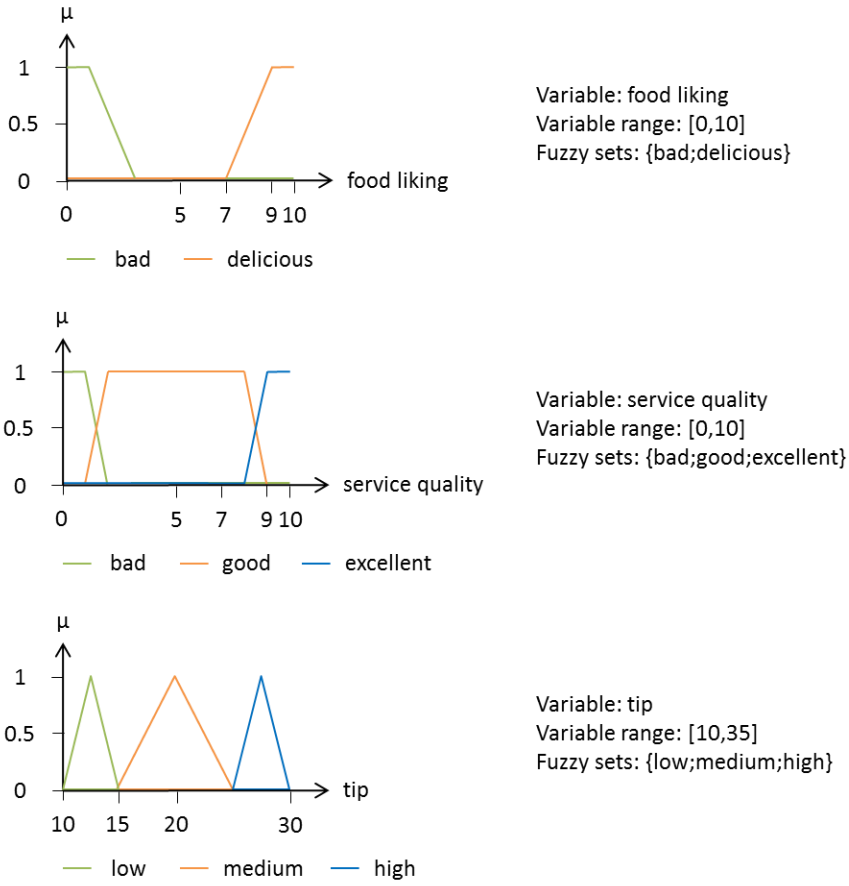


Figure 2.5: Fuzzy functions of the variables food liking, service quality and tip (adapted from Dernoncourt, 2011).

b. Fuzzy rules

As the choice of membership function's shape, the definition of the operators in fuzzy logic is chosen according to the expertise modelled. Operators include intersection (AND), union (OR) and complement (NOT). Triangular norm (t-norm) and conorm (t-conorm) are operations corresponding to logical conjunction (i.e. intersection) and logical disjunction (i.e. union) in fuzzy logic systems. T-norm and conorm are used to combine criteria in multi-criteria decision making. Zadeh's definition of intersection used the Gödel t-norm, but other definitions are possible using different t-norms. Common examples are Product t-norm or

Lukasiewicz t-norm (Hudelot, 2014). In Table 2.4, the operators are presented according to Gödel's and Product's semantic.

Table 2.4: Fuzzy operators. A and B are two fuzzy sets. Intersection (AND) is defined as $\mu_{A\cap B}(x)$, union (OR) as $\mu_{A\cup B}(x)$ and complement (NOT) as $\mu_{\bar{A}}(x)$.

Operators	AND	OR	NOT
Zadeh	$\min(\mu_A(x), \mu_B(x)$	$\max(\mu_A(x), \mu_B(x))$	$1 - \mu_A(x)$
Product	$\mu_A(x) \times \mu_B(x)$	$\mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x)$	$1 - \mu_A(x)$

The operators allow to construct fuzzy rules such as "If $x \in A$ AND $y \in B$ then $z \in C$", with A, B, and C as fuzzy sets. By definition all the rules of a rule basis are activated at the same time and linked by an OR operator.

To predict the tip given in a restaurant, three rules are created. These rules represent the knowledge-based expressing and predicting the state of the system.

Rule 1: If (food is bad) and (service is bad) then (tip will be low)

Rule 2: If (service is good) then (tip will be medium)

Rule 3: If (food is delicious) and (service is excellent) then (tip will be high)

For the application of this example, we suppose the input values equal to 8 for the food liking and 8.5 for the service quality. For each of the three fuzzy rules, we obtain an output (Figure 2.6). For Rule 1 and Rule 3, operators are combined with a Product t-norm and we obtained respectively 0 (0×0) and 0.25 (0.5×0.5). For Rule 2, there is no combination because only one variable is considered.

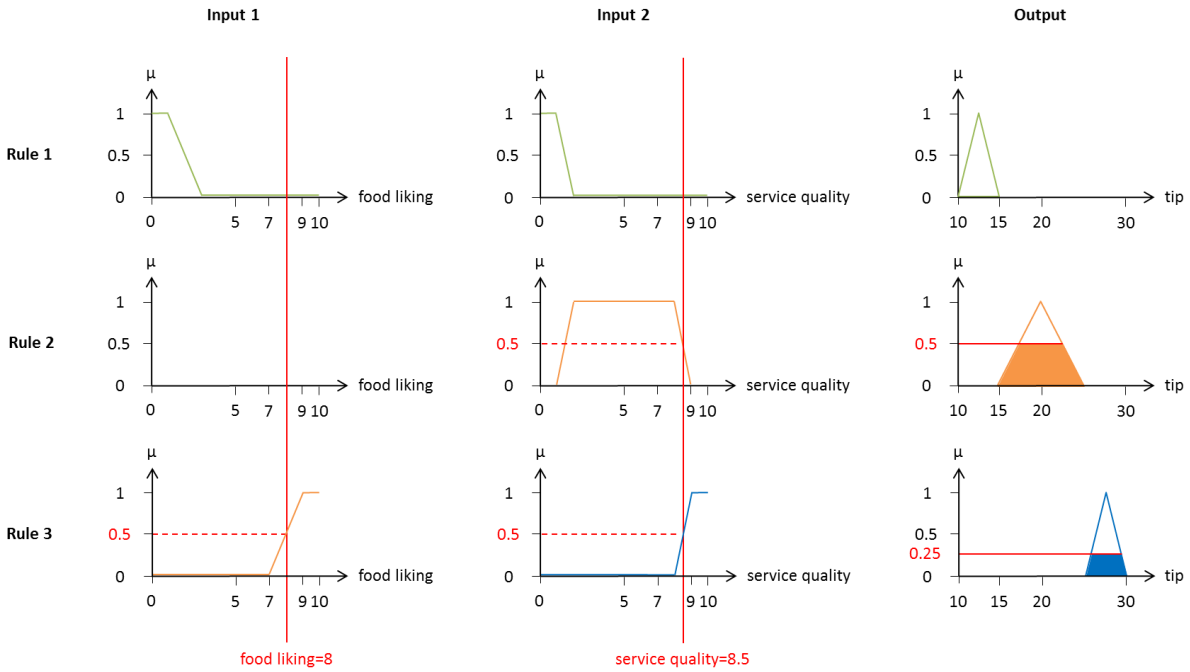


Figure 2.6: Fuzzy rules. Inputs are food liking and service quality, both scored on a 10-point scale. Output is the percentage of tip given. Product t-norm is used to combine both inputs.

Regarding the outputs, the remaining question is what will be the tip percentage, knowing that the food liking is rated 8 on 10 and the quality of the service 8.5 out of 10. The two activated rules (Rule 2 and Rule 3) are combined together using the OR operator. It leads to a fuzzy set output coupling both rules (trapezoids blues and orange of the Figure 2.7). In some applications where decision or prediction is needed, a following step, called defuzzification, translate the aggregated fuzzy set into a numerical result.

c. Defuzzification

As with the fuzzy membership functions and fuzzy operators, the fuzzy system designer must choose from several possible defuzzification methods. In Figure 2.7 the results are presented following the average maxima method. This defuzzification defines the output (decision of the percentage of the tip) as being the average of the abscissas of the maxima of the fuzzy set resulting from the aggregation of the outputs. In this case, the maxima are respectively 25 and 30, thus the percentage of tip given will of 27.5 %.

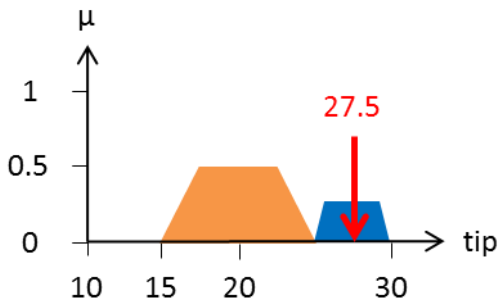


Figure 2.7: Defuzzification of the outputs.

The designer of a fuzzy system has to make a number of important choices (membership functions' shape, operators, defuzzification method) which impact the final results. These choices are mainly based on the expert's advice or on statistical analysis of previous data, in particular to define the membership functions and the fuzzy rules.

Thus, the power of fuzzy logic is to make possible the establishment of inference systems (choices of operators) whose decisions are seamless, flexible and non-linear, closer to human behavior than conventional logic. In addition, the rules are expressed in natural language. This has many advantages, such as including the knowledge of a non-computer expert at the heart of a decision-making system, or modelling more finely some aspects of natural language.

VI. Conclusion

The research question at the core of this thesis was tackled through a series of methods from several scientific disciplines: analytical chemistry, sensory evaluation, knowledge integration, modelling and optimization. In the three following chapters, the three main studies achieved during the thesis are presented. The 16 wines, characterized through analytical and sensory analysis (Villière et al., **Data paper included in this manuscript**), constitute the sample set. These data were used as input data and allowed the testing of modelling approaches on complex real food products.

In Chapter 3, a predictive modelling strategy based on molecular structure of odorants is presented (Roche et al., **Article 1 of this manuscript**). The aim was to predict complex mixtures perceptual similarity from odorants' structural parameters obtained owing to the Dragon software (Talete, Milan, Italy). The model, once optimized (Roche et al., **Article 1 of this manuscript**), was applied to the 16 wines. The modelling performance was assessed through the correlation between perceptual wine odor similarity predicted by the model and similarity inferred from sensory profiling data.

In Chapter 4, the construction of an ontology for odor perceptual space (OOPS) is presented (Roche et al., **Article 2 of this manuscript**). For the 49 odorants identified in the wines, we collected their odor descriptors (OD) from different databases. Owing to the flavorists' expertise collected, we linked the corpus of OD to the corpus relying on expertise, namely odor qualities (OQ) and to the 15 odor sensory attributes (OSA) used by the sensory panel to perform the wines profiling. The ontology was implemented in the dedicated Protégé® software and was used as a prediction tool to predict the sensory profiles of two wines of the sample set on the basis of their molecular composition.

Chapter 5 presents a broader predictive strategy, which combines ontology, fuzzy logic and optimization methodologies to form an operational expert system (Roche et al., **Article 3 of this manuscript**). Input values of the model were the 49 odorants identified in the 16 wines further translated in OD and OQ thanks to the ontology previously developed (Chapter 4). Fuzzy logic was applied for its ability to cope with non-linear systems expressed by expert flavorists, i.e. the combination of OQ contributing to the perception of OSA. The fuzzy rules were determined through data-driven optimization. The developed model was able to predict the odor profile of wines through the estimation of the intensity of OSA. Predictions were compared with the odor profile obtained by sensory evaluation.

FIRST AXIS:

MOLECULAR STRUCTURE APPROACH

CHAPTER 3: STRUCTURAL APPROACH

I. Introduction

Olfactory research was and is still challenged by predicting odor characteristics of odorants on the basis of their molecular structure. Different teams have been working on establishing successful predictive approaches, but many of them apply to single odorants. However, odors we perceive in every-day life are mixtures including many different odorants at varying concentrations. The odor quality of such mixtures can be perceived as elemental (components odors can be perceived within the mixture), configural (components odors blend into a new odor perceived as an entity) or partially elemental and/or configural (Thomas-Danguin et al., 2014).

An attempt to predict odor characteristic of mixtures of odorants was performed by Snitz et al. (2013). They developed an angle distance model which successfully predicts the similarity of complex mixtures composed of iso-intense components, on the basis of their molecular structure. In this chapter, we present how we tested this model using a dataset of elemental and configural mixtures made of 6 odorants, their sub-mixtures, and odorants alone, for which perceptual similarity data were obtained by a panel of 60 subjects for 63 pair comparisons (Romagny et al., 2018). Moreover, we present how we upgraded this model to take into account the intensity of each odorant in mixtures to be able to account for real odor mixtures such as the 16 wines considered in the thesis.

The application of the Snitz' model to the elemental and configural mixtures made of 6 odorants, their sub-mixtures and odorants alone is presented in the **ARTICLE 1** entitled "Predicting odor similarity of complex mixtures from molecular structure" (Roche, A, Thomas-Danguin, T, Perrot, N & Mainland, J, in preparation). The application of the upgraded model to the set of wines is presented in the second part of the chapter.

II. ARTICLE 1

Predicting odor similarity of complex mixtures from molecular structure

Alice Roche[1,2], Thierry Thomas-Danguin[1], Nathalie Perrot[3], Joel Mainland[2,4]

1. CSGA, UMR1324 INRA, UMR6265 CNRS, AgroSup Dijon, Université de Bourgogne Franche - Comté, Dijon, France
2. Monell Chemical Senses Center, Philadelphia, PA, USA
3. GMPA, UMR782 INRA, AgroParisTech, Université Paris Saclay, Thiverval-Grignon, France
4. Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA

* Correspondence to be sent to: Thierry Thomas-Danguin thierry.thomas-danguin@inra.fr or Joel Mainland jmainland@monell.org

This manuscript is in preparation.

Keywords: angle distance, molecular structure, odorants mixture, similarity, odor prediction

Abstract

Predicting perceptual characteristics of odor stimuli on the basis of their molecular structure is a challenging task. To date, most approaches were applied to monomolecular odor stimuli. However, odors we encountered in our day-to-day lives result from the perception of complex mixtures of odorants. In this study, we applied the angle distance model, previously proposed to predict the similarity between mixtures of odorants on the basis of their structural parameters, to mixtures made of 6 odorants, their sub-mixtures and single odorants (63 pair comparisons). We extended the model to take into account the relative intensity of each odorant within the mixtures. The ratio-weighted angle distance model well predicted the perceptual similarity ($r = 0.75$, $p < 0.001$) between the odor stimuli of our experimental dataset that included mixtures with different levels of complexity and odorants at various intensities. Moreover, the model was found to be able to well predict perceptual similarity for mixtures processed in a more elemental or configural way by the human olfactory system. Therefore, the ratio-weighted angle distance model, modified to account for mixtures' components intensity, constitutes a step forward to the science of smell since it is efficient to predict similarity between complex odor mixtures, and in a way consistent with the processing of odor objects by the human brain.

Introduction

Olfactory research was and is still challenged by predicting odor characteristics of odor stimuli such as their detection threshold (Abraham et al., 2012), pleasantness (Khan et al., 2007; Kermen et al., 2011), intensity (Edwars & Jurs, 1989; Chastrette et al., 1998; Thomas-Danguin & Chastrette, 2002) or their smell quality (Wise et al., 2000; Keller et al., 2017). Several predictive models embedded structural parameters of molecules in their approach. The number of structural parameters to be considered can rise to thousands thanks to recent progress in software (Dragon software, Talete, Milan, Italy) and represent a very large set of parameters of molecules, varying from simple functional group counts to normalized eigenvalue sums of the powers of connectivity matrices.

Focusing on predicting the smell of molecules, approaches based on the structure-odor-relationships (SOR), identified many specific rules linking structure to odor, i.e. what structure may provide a specific odor note (e.g. ‘‘musk’’ in Lavine et al., 2012), but failed to

produce a general framework for measuring smell (Rossiter, 1996). Indeed, molecules that have a close structure can elicit very different odors: carvone (−) (CAS 6485-40-1) smells minty, whereas its enantiomer (+) (CAS 2244-16-8) smells like caraway. In contrast, molecules that have a different structure can elicit similar odors: muscone (CAS 541-91-3) and musk ketone (CAS 81-14-1) both smell musky. With the aim to overcome this issue, machine learning approaches are emerging to construct systematic models in order to better characterize odorants on the basis of their molecular structure.

Kumar et al. (2015) showed that it was possible to use selective structural parameters to predict the perceptual qualities of molecules (e.g. “butter”, “fruity”, “leather”…) by designing a random forest classifier. Following works kept developing machine learning algorithms to predict odor characteristics of molecules based on their structural parameters. For instance, the best-performing models from the DREAM Olfaction project (Keller et al., 2017) were able to accurately predict 8 out of 19 semantic descriptors “garlic,” “fish,” “sweet,” “fruit,” “burnt,” “spices,” “flower,” and “sour” ($r>0.5$) of a dedicated dataset of molecules (Keller & Vosshall, 2016).

Although these predictive approaches were quite successful, only a few studies attempted to predict odor quality, and all applied to single molecules. If basic knowledge on monomolecular stimuli is essential, it does not reflect the ecological functioning of the sense of smell, which has to deal with complex mixtures of odorants at varying concentrations. To efficiently treat complex mixtures, it has been shown that the olfactory system can process odor mixture through a configural process in which components’ odors blend into a new odor perceived as a distinct odor object or through an elemental process in which components’ odors can be perceived within the mixture (Stevenson & Wilson, 2007; Thomas-Danguin et al., 2014). To our knowledge, the only predictive work dealing with odor quality of odorants’ mixtures was performed by Snitz et al. in 2013, who developed a model to predict the perceptual similarity of multi-molecular mixtures from the molecular structure of the components. In this approach, mixtures were represented as single structural vectors to take into account the mixture as a whole and not only as a sum of its constituents. Interestingly, such approach is consistent with a configural brain processing of complex odor mixtures (Gottfried, 2010; Thomas-Danguin et al., 2014). The algorithm, which considered a series of structural parameters of the mixed odorants, was found to provide consistent correlations between predicted and actual perceptual similarity ($r = 0.49$, $p < 0.001$). An optimized version of this model yielded a correlation of $r = 0.85$ ($p < 0.001$) between predicted and actual

perceptual similarity. Although applicable to mixtures of odorants, the main limitation of the model was to consider odorants only at the same intensity. Such iso-intense mixtures were experimentally created but do not occur in nature.

In order to overcome this limitation, we developed an extended version of the angle distance model proposed by Snitz et al. (2013) that was able to account for non iso-intense mixtures, i.e. the ratio-weighted angle distance model. We tested the ratio-weighted angle distance model with a set of mixtures made of up to 6 odorants (Romagny et al., 2018), among which two mixtures of the same 6 odorants but in different proportions were known to be processed respectively configurally and elementally (Sinding et al., 2013). The dataset comprised similarity between these mixtures, sub-mixtures and single odorants corresponding to a total of 63 pair comparisons (Romagny et al., 2018; Table S1).

Materials and methods

Dataset of odor stimuli similarity

The dataset of similarity between odor stimuli was from the study by Romagny et al. (2018). It included 63 similarity values corresponding to 63 comparisons between two stimuli including mixtures and single components. One mixture, the RC configural mixture (RC-conf) contained 6 odorants (Table 1) at specific concentration and has been shown to elicit configural perception and to evoke the specific odor of grenadine (Red Cordial; Sinding et al. 2013). The RC elemental mixture (RC-elem) contained the same 6 odorants but at different ratio and did not elicit the Red Cordial configural odor. The other stimuli in the dataset were 30 sub-mixtures and the 6 single odorants. The sub-mixtures were coded with the acronym of the odorants in the stimuli and their ratio (either RC-conf or RC-elem). Sub-mixtures at the RC-conf ratio corresponded to Sub-con mixtures and the ones at RC-elem ratio corresponded to Sub-elem mixtures. For instance, the stimuli "BEA_RC-conf" corresponded to the sub-mixture of β-ionone and Ethyl acetate at the RC-conf concentration levels and the stimuli "IAVFDB_RC-elem" corresponded to the sub-mixture of Isoamyl acetate, Vanillin, Frambinone, Damascenone and β-ionone at the RC-elem concentration levels. The summary of all the stimuli studied in Romagny et al. (2018) is shown in Table S1.

Table 1: The 6 odorants used in the mixture dataset.

Odorant	Acronym	CAS	Canonical SMILES
Vanillin	V	80-14-4	COC1=C(C=CC(=C1)C=O)O
Frambinone	F	5471-51-2	CC(=O)CCC1=CC=C(C=C1)O
Isoamyl acetate	IA	123-92-2	CC(C)CCOC(=O)C
Damascenone	D	23696-85-7	CC=CC(=O)C1=C(C=CCC1(C)C)C
β-ionone	B	79-77-6	CC1=C(C(CCC1)(C)C)C=CC(=O)C
Ethyl acetate	EA	141-78-6	CCOC(=O)C

The 6-component mixtures, sub-mixtures and single odorants were rated on their similarity with the RC-conf or RC-elem 6-component mixtures. The similarity data were obtained by direct rating using a linear scale in pair comparison between the odor of the reference RC-conf or RC-elem mixture and the odor of a sample including either a single odorant, a sub-mixture or the RC-conf or RC-elem mixtures (controls). A total of 63 pair comparisons were evaluated in four separate experiments as described in Romagny et al. (2018). Ratings were then converted into similarity scores ranging from 0 to 10 (Table S2).

For each pair comparison we also calculated the overlap between the two stimuli following the Equation 1, *n* being the number of shared components between the two stimuli and N being the number of components in the largest mixture (6 components).

Equation 1: $$Overlap = \frac{n}{N}$$

Molecular structure of odorants

The 6 odorants studied in this paper were described by 4870 structural parameters using Dragon software (v.6, Talete, Milan, Italy). In Table 2A, the odorants are described by the first 8 structural parameters obtained from the software.

Table 2: The 6 odorants described by the first 8 structural parameters (A) from the Dragon software as raw data and (B) as normalized data. MW: molecular weight, AMW: average molecular weight, Sv: sum of atomic van der Waals volumes, Se: sum of atomic Sanderson electronegativities, Sp: sum of atomic polarizabilities, Si: sum of first ionization potentials, Mv: mean atomic van der Waals volume, Me: mean atomic Sanderson electronegativity.

A

Odorant	MW	AMW	Sv	Se	Sp	Si	Mv	Me
V	152.16	8.01	12.25	19.52	12.41	21.29	0.65	1.03
F	164.22	6.84	14.59	23.96	15.48	26.91	0.61	1
IA	130.21	5.66	12.12	22.84	13.24	26.33	0.53	0.99
D	190.31	5.95	18.46	31.28	20.31	35.95	0.58	0.98
B	192.33	5.66	18.98	33.16	21.07	38.36	0.56	0.98
EA	88.12	6.29	7.54	14.19	7.96	16.08	0.54	1.01

B

Odorant	MW	AMW	Sv	Se	Sp	Si	Mv	Me
V	0.61	1	0.41	0.28	0.34	0.23	1	1
F	0.73	0.5	0.62	0.51	0.57	0.49	0.69	0.44
IA	0.4	0	0.4	0.46	0.4	0.46	0	0.35
D	0.98	0.12	0.95	0.9	0.94	0.89	0.42	0.04
B	1	0	1	1	1	1	0.26	0
EA	0	0.27	0	0	0	0	0.09	0.75

Since the different structural parameters measure properties on differing scales we normalized the Dragon data. For each structural parameter we had a set of 6 values (*sv*) (e.g. 1.03, 1, 0.99, 0.98, 0.98, 1.01 for the structural parameter Me). Each values *v* in the list *sv* was normalized to the value *vn* by the Equation 2.

Equation 2:

$$vn = \frac{v - \min(sv)}{\max(sv) - \min(sv)}$$

We ended up with values for each structural parameter ranging between 0 and 1. Each component was then described by a structural vector of 4870 dimensions. The normalized values of the Table 2A are presented in Table 2B.

The angle distance model

Snitz et al. (2013) developed the angle distance model that allowed predicting odorant-mixture perceptual similarity from odorant-mixture structure. It consisted in summing the structural vectors of its components and dividing it by its norm to eliminate the effect of the number of components in a mixture on the size of the mixture vector. Then the distance between two mixtures vector (mixture U and mixture V) is defined as the angle between the two vectors, given by Equation 3 where the dot product between the vectors is $\vec{U} \cdot \vec{V}$ and the norms of the vectors are $|\vec{U}||\vec{V}|$.

Equation 3: $$\theta(\vec{U},\vec{V}) = \cos^{-1}\left(\frac{\vec{U}\cdot\vec{V}}{|\vec{U}||\vec{V}|}\right)$$

This angle distance was further developed in an optimized version. In this case, the dimension of the structural vectors was reduced to 21 structural parameters (Table 4). We considered both the angle distance model (non-optimized) and the optimized angle distance model. Table S3 contains the odorants we modeled and their structural parameters values for the optimized angle distance model.

Table 4: List of the 21 structural parameters for the optimized angle distance model. Listed are the names, indices and a brief definition of the 21 parameters, named molecular descriptors within the Dragon software. The Dragon descriptors SpMin3_Bh(v), SM02_EA(ed), SM03_EA(dm), SM10_EA(dm), SM13_EA(dm) and Eig05_EA(ed) were abbreviated respectively BELv3, ESpm02x, ESpm03d, ESpm10d, ESpm13d and EEig05x in Snitz et al. (2013) due to different software versions.

Number	Index out of the 4870 descriptors	Abbreviation	Description
1	45	nCIR	Number of circuits (ring descriptors)
2	76	ZM1	First Zagreb index (topological indices)
3	97	GNar	Narumi geometric topological index (topological indices)
4	122	S1K	1-path Kier alpha-modified shape index (topological indices)
5	187	piPC08	Molecular multiple path count of order 8 (walk and path counts)
6	936	MATS1v	Moran autocorrelation of lag 1 weighted by van der Waals volume (2D autocorrelations)
7	942	MATS7v	Moran autocorrelation of lag 7 weighted by van der Waals volume (2D autocorrelations)
8	984	GATS1v	Geary autocorrelation of lag 1 weighted by van der Waals volume (2D autocorrelations)
9	1103	SpMin3_Bh(v)	Smallest eigenvalue n. 3 of Burden matrix weighted by van der Waals volume (Burden eigenvalues)
10	1286	SM02_EA(ed)	Spectral moment of order 2 from edge adjacency mat. weighted by edge degree (edge adjacency indices)
11	1315	SM03_EA(dm)	Spectral moment of order 3 from edge adjacency mat. weighted by dipole moment (edge adjacency indices)
12	1322	SM10_EA(dm)	Spectral moment of order 10 from edge adjacency mat. weighted by dipole moment (edge adjacency indices)
13	1325	SM13_EA(dm)	Spectral moment of order 13 from edge adjacency mat. weighted by dipole moment (edge adjacency indices)
14	1417	Eig05_EA(ed)	Eigenvalue n. 5 from edge adjacency mat. weighted by edge degree (edge adjacency indices)
15	1806	RDF035v	Radial Distribution Function - 035 / weighted by van der Waals volume (RDF descriptors)
16	2191	G1m	1st component symmetry directional WHIM index / weighted by mass (WHIM descriptors)
17	2202	G1v	1st component symmetry directional WHIM index / weighted by van der Waals volume (WHIM descriptors)
18	2213	G1e	1st component symmetry directional WHIM index / weighted by Sanderson electronegativity (WHIM descriptors)
19	2248	G3s	3rd component symmetry directional WHIM index / weighted by I-state (WHIM descriptors)
20	2452	R8u+	R maximal autocorrelation of lag 8 / unweighted (GETAWAY descriptors)
21	2646	nRCOSR	number of thioesters (aliphatic) (Functional group counts)

The angle distance model previously described (non-optimized and optimized versions) considered components at the same intensity. In the present study we upgraded the model to account for non iso-intense mixtures. To do so, we weighted each structural vector by the ratio of the corresponding odorant in the mixture.

Data analysis

Simple linear regressions were performed to model the relationships between two variables: variable 1 being the perceptual similarity of the stimuli and the second variable being the overlap of the stimuli (i.e. the number of shared components in both stimuli) or the angle distance of the stimuli. In addition, the Pearson correlation (r) was calculated as a measure of the linear correlation between the two variables.

In order to compare the different correlations resulting from the different versions of the angle distance model, we estimated the statistical differences between two dependent correlations owing to the Steiger's test (Steiger, 1980) following the function *cordif.dep* of the R "multilevel" package (Bliese, 2016).

Results and discussion

Perceptual similarity and molecular overlap

First of all, we evaluated the relationship the perceived similarity and the molecular overlap between stimuli (6-components mixtures, their sub-mixtures and single odorants). We found a significant and rather high correlation ($r = 0.65$, $p < 0.001$; Figure 1), a result with those of Bushdid et al. (2014) suggested that similarity among mixtures increased with the number of shared components between the mixtures. However, when we considered only the red dots in Figure 1, which corresponded to either the similarity between the configural mixture (RC-conf) and the elemental mixture (RC-elem) shown as dark red dots or the similarity between sub-mixtures of the configural mixture (Sub-conf) and the elemental mixture (RC-elem) shown as light red dots, then the correlation dropped and was no more significant ($r = 0.1$, $p = 0.73$). This observation was important because the mixtures RC-conf and RC-elem were

composed of the same 6 molecules, only the ratio between the odorants was different, so their overlap is 100% but their perceptual similarity is low (3.66). With regard to sub-mixtures the overlap decreased as the number of components decreased, but the similarity was not at all related to this molecular overlap (no correlation). In contrast, considering the blue and green dots in Figure 1, especially for molecular overlap higher than 0.5, it can be seen that high similarity values were maintained, which underlined that even if similarity among mixtures increased with the number of shared components, for the most complex mixtures, the ratio between odorants was critical and likely drives the perceptual similarity. Therefore it became obvious that the integration of relative intensity of the components in a mixture is essential in odor quality prediction approaches since it can reflect the component ratio, a feature largely used by the olfactory system and the brain to code and process complex mixtures of odorants.

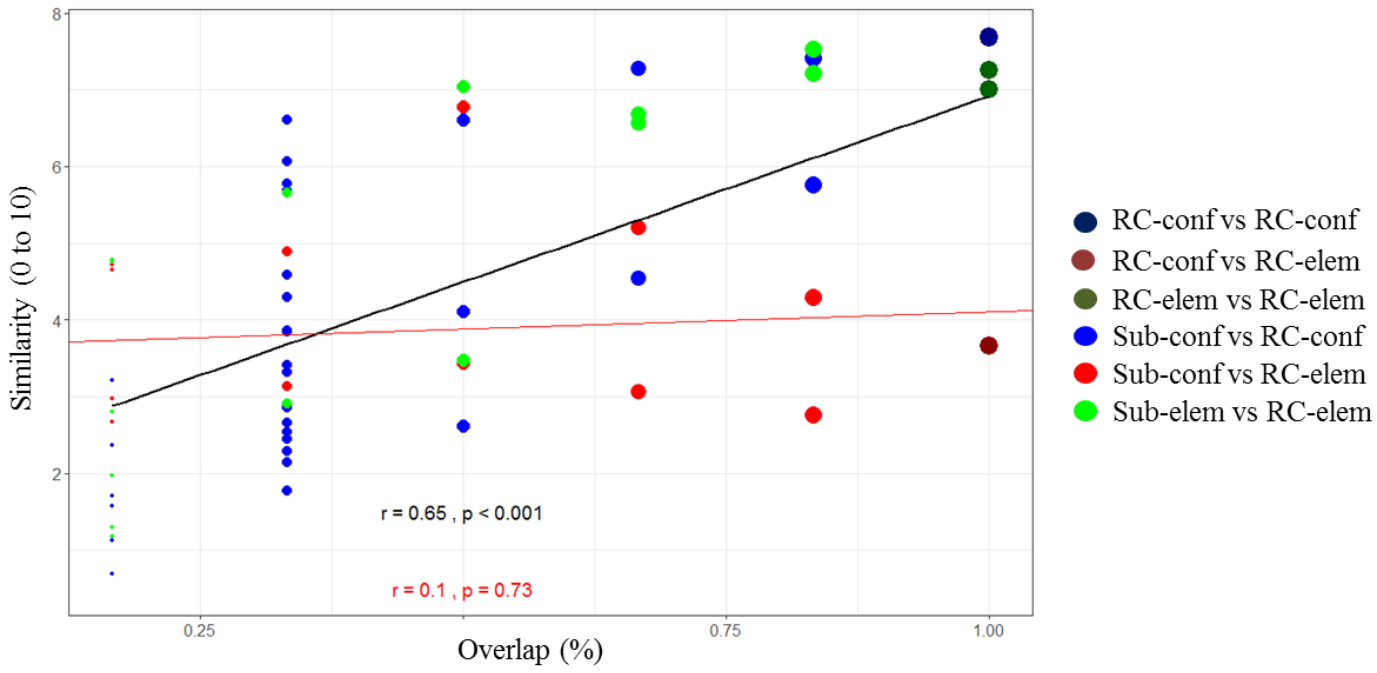


Figure 1: Correlation between perceived similarity and molecular overlap of 63 pair of complex odor stimuli. Each dot reflected a comparison between two stimuli (6-component mixture RC-conf or RC-elem vs mixtures, sub-mixtures or single molecules). A linear regression was performed on the entire dataset (black line) and another one only considering the red shaded dots (red line). The size of the dots varied according to the number of molecules in the stimuli (from 1 to 6).

Prediction of perceptual similarity

We applied the angle distance model following the methodology described in Snitz et al. (2013) first in its non-optimized version, i.e. by taking into account the 4870 structural parameters provided by the Dragon software (Figure 2A). The correlation between the perceived similarity and the angle distance for the 63 pair stimuli was good ($r = -0.62$, $p < 0.001$). We also applied the optimized version of the angle distance model by taking into account only 21 structural parameters among the 4870 provided by the Dragon software. The optimized angle distance model led to a slight increase of the correlation coefficient ($r = -0.63$, $p < 0.001$; Figure 2B). The slight correlation increase was not significant ($p = 0.75$ from the Steiger's test), but since the optimized angle distance model was based on a lower number of structural parameters, it can be considered better than the original angle distance model. In this case, the 21 structural parameters selected by Snitz et al. (2013) to optimize their model well fitted to our dataset.

We upgraded the angle distance model to take into account the difference in ratio between the components of the mixtures included in our dataset. Indeed, the original version (Snitz et al., 2013) applied only to iso-intense mixtures. To do so, we weighted each structural vector by the ratio of the odorants included in the mixture. We obtained a ratio-weighted angle distance model, which was able to account for the ratio of each molecule within the mixtures. The ratio-weighted angle distance model, which used the 4870 structural parameters provided by the Dragon software, led to better correlation results ($r = -0.75$, $p < 0.001$; Figure 2C) compared to the angle distance model. The correlation increase was proven significant owing to the Steiger's test ($p = 0.03$). Moreover, the ratio-weighted angle distance model pretty well succeed in predicting the similarity between the configural mixture (RC-conf) and the elemental mixture (RC-elem, dark red dot in Figure 2), which contained the same 6 odorants but in a different ratio. On the whole, the ratio-weighted angle distance model better predicted the similarity between sub-mixtures of the configural mixture (Sub-conf) and the elemental mixture (RC-elem, light red dots in Figure 2), as well as the similarity between sub-mixtures of the elemental mixture (Sub-elem) and the elemental mixture (RC-elem, light green dots in Figure 2).

We then used only the 21 structural parameters from the optimized angle distance model (Snitz et al., 2013) but taking into account the ratio-weighted approach. However, this optimized ratio-weighted angle distance model led to less good correlation with perceptual

similarity (r = -0.69, p < 0.001; Figure 2D), confirmed by the Steiger's test (p = 0.005). In this case, the 21 structural parameters selected by Snitz et al. (2013) did not fit to our dataset. Further work required to optimize the structural parameters to cope with such ratio-weighted angle distance model.

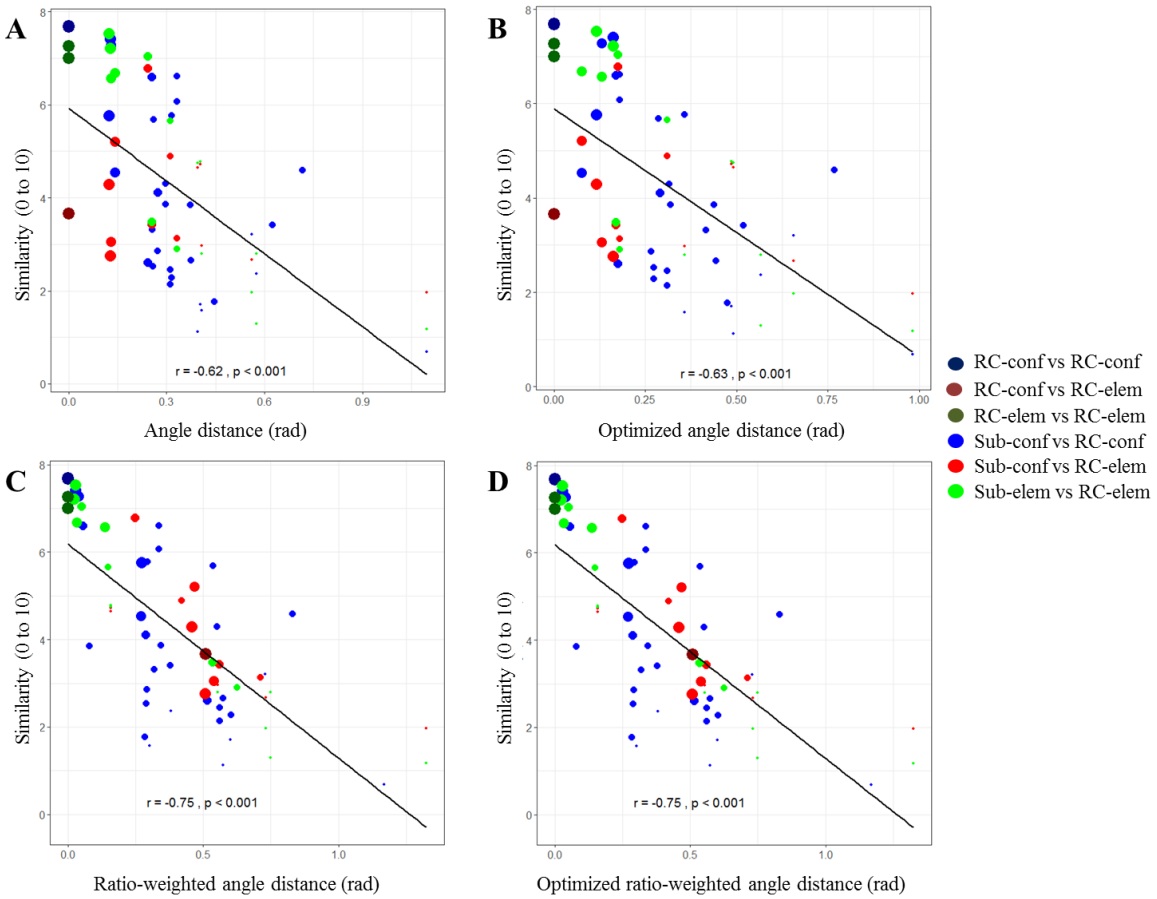


Figure 2: Performance of the different versions of the angle distance model. Each dot reflected a comparison between two stimuli (6-component mixture RC-conf or RC-elem vs mixtures, sub-mixtures or single molecules). The size of the dots varied according to the number of molecules in the stimuli (from 1 to 6). A linear regression was performed (black line) to evaluate the prediction performances of (A) the angle distance model, (B) the optimized angle distance model, (C) the ratio-weighted angle distance model and (D) the optimized ratio-weighted angle distance model.

Conclusion

We applied the angle distance model, previously proposed to predict the similarity between iso-intense mixtures of odorants on the basis of their structural parameters, to a dataset of similarity values obtained with mixtures made of 6 odorants at a different ratio, their sub-mixtures and single odorants. One advantage of the angle distance model was to consider odorants mixture as a whole rather by constructing molecular structure vector for each mixture studied. This approach was consistent with object-oriented processing of complex olfactory mixtures. We demonstrated the need to extend the model to take into account the relative intensity of each molecule within the mixtures. The upgraded version of the model, the ratio-weighted angle distance model, better predicted the perceptual similarity between the odor stimuli of our experimental dataset. The upgraded model was found to be able to correctly predict the odor similarity for elmental or configural mixtures from the structural parameters of the components and their ratio. Therefore, the angle distance model, modified to account for mixtures' components intensity, constitutes a step forward in the science of smell since it is efficient to predict similarity between complex odor mixtures, and in a way consistent with the processing of odor objects by the human brain. One limitation of our work was to use of concentration ratio of odorants in mixture instead of their intensity. Therefore, further work is needed to gain prediction power using intensity ratio for instance by integrating models of the non-linear dose-response curve of odorants (Chastrette et al., 1998).

Acknowledgments

This work was supported by grants from the Conseil Régional Bourgogne, Franche-Comté. AR was supported by an INRA thesis fellowship.

References

Abraham, MH, Sánchez-Moreno, R, Cometto-Muñiz, JE & Cain, WS. 2012. An algorithm for 353 odor detection thresholds in humans. Chemical senses, 37(3):207-218. doi:10.1093/chemse/bjr094.

Bliese, P. 2016. multilevel: Multilevel Functions. R package version 2.6. https://CRAN.R-project.org/package=multilevel

Bushdid, C, Magnasco, MO, Vosshall, LB & Keller, A. 2014. Humans can discriminate more than 1 trillion olfactory stimuli. Science, 343(6177):1370-1372. doi:10.1126/science.1249168.

Chastrette, M, Thomas-Danguin, T & Rallet, E. 1998. Modelling the human olfactory stimulus-response function. Chemical senses, 23(2):181-196. doi:10.1093/chemse/23.2.181.

Edwards, PA & Jurs, PC. 1989. Correlation of odor intensities with structural properties of odorants. Chemical senses, 14(2):281-291. doi:10.1093/chemse/14.2.281.

Gottfried, JA. 2010. Central mechanisms of odour object perception. Nature reviews neuroscience, 11:628-641. doi:10.1038/nrn2883.

Keller, A & Vosshall, LB. 2016. Olfactory perception of chemically diverse molecules. BMC Neuroscience, 17:55. doi:10.1186/s12868-016-0287-2.

Keller, A, Gerkin, RC, Guan, Y, Dhurandhar, A, Turu, G, Szalai, B, Mainland, JD, Ihara, Y, Yu, CV, Wolfinger, R, Vens, C, Schietgat, L, De Grave, K, Norel, R, DREAM Olfaction Prediction Consortium, Stolovitzky, G, Cecchi, GA, Vosshall, LB & Meyer, P. 2017. Predicting human olfactory perception from chemical features of odor molecules. Science, 355(6327):820-826. doi:10.1126/science.aal2014.

Kermen, F, Chakirian, A, Sezille, C, Joussain, P, Le Goff, G, Ziessel, A, Chastrette, M, Mandairon, N, Didier, A, Rouby, C & Bensafi, M. 2011. Molecular complexity determines the number of olfactory notes and the pleasantness of smells. Scientific reports, 1:206. doi:10.1038/srep00206.

Khan, RM, Luk, C-H, Flinker, A, Aggarwal, A, Lapid, H, Haddad, R & Sobel, N. 2007. Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. Journal of neuroscience, 27(37):10015-10023. doi:10.1523/JNEUROSCI.1158-07.2007.

Kumar, R, Kaur, R, Auffarth, B & Bhondekar, AP. 2015. Understanding the odour spaces: A step towards solving olfactory stimulus-percept problem. Plos one, 10(10): e0141263. doi:10.1371/journal.pone.0141263.

Lavine, BK, White, C, Mirjankar, N, Sundling, CM & Breneman, CM. 2012. Odor-structure relationship studies of Tetralin and Indian musks. Chemical senses, 37(8):726-736. doi:10.1093/chemse/bjs058.

Romagny, S, Coureaud, G & Thomas-Danguin, T. 2018. Key odorants or key associations? Insights into elemental and configural odour processing. Flavour and fragrance journal, 33:97-105. doi:10.1002/ffj.3429.

Rossiter, KJ. 1996. Structure-odor relationships. Chemical reviews, 96(8):3201-3240. doi:10.1021/cr950068a.

Sinding, C, Thomas-Danguin, T, Chambault, A, Béno, N, Dosne, T, Chabanet, C, Schaal, B & Coureaud, G. 2013. Rabbit neonates and human adults perceive a blending 6-component odor mixture in a comparable manner. Plos one, 8(1):e53534. doi:10.1371/journal.pone.0053534.

Snitz, K. 2016. Describing odorspace. P63. In: XXVIth Annual meeting of the european chemoreception research organization, ECRO 2016: Athens, Greece, 7-10 September, 2016, Organizer: Marika Kapsimali. doi:10.1093/chemse/bjw120.

Snitz, K, Yablonka, A, Weiss, T, Frumin, I, Khan, RM & Sobel, N. 2013. Predicting odor perceptual similarity from odor structure. Plos computational biology, 9(9):e1003184. doi:10.1371/journal.pcbi.1003184.

Steiger, JH. 1980. Tests for comparing elements of a correlation matrix. Psychological bulletin, 87(2):245-251. doi:10.1037/0033-2909.87.2.245.

Stevenson, RJ & Wilson, DA. 2007. Odour perception: An object-recognition approach. Perception, 36(12):1821-1833. doi:10.1068/p5563.

Thomas-Danguin, T & Chastrette, M. 2002. Odour intensity of binary mixtures of odorous compounds. Comptes rendus biologies, 325(7):767-772. doi:10.1016/S1631-0691(02)01485-3.

Thomas-Danguin, T, Sinding, C, Romagny, S, El Mountassir, F, Atanasova, B, Le Berre, E, Le Bon, A-M & Coureaud, G. 2014. The perception of odor objects in everyday life: A review on the processing of odor mixtures. Frontiers in psychology, 5:504. doi:10.3389/fpsyg.2014.00504.

Wise, PM, Olsson, MJ & Cain, WS. 2000. Quantification of odor quality. Chemical senses, 25(4):429-443. doi:10.1093/chemse/25.4.429.

Supplementary data

Table S1: Composition of the stimuli. The stimuli are described by their number of components and their ratio of components either at the RC-conf ratio (+) or at the RC-elem ratio (o).

Stimulus	Number of components	V	F	IA	D	B	EA
B_RC-conf	1					+	
B_RC-elem	1					o	
D_RC-conf	1				+		
D_RC-elem	1				o		
EA_RC-conf	1						+
EA_RC-elem	1						o
F_RC-conf	1		+				
F_RC-elem	1		o				
IA_RC-conf	1			+			
IA_RC-elem	1			o			
V_RC-conf	1	+					
V_RC-elem	1	o					
BEA_RC-conf	2					+	+
DB_RC-conf	2				+	+	
DEA_RC-conf	2				+		+
EAB_RC-conf	2					+	+
EAB_RC-elem	2					o	o
FB_RC-conf	2		+			+	
FD_RC-conf	2		+		+		
FEA_RC-conf	2		+				+
IAB_RC-conf	2			+		+	
IAD_RC-conf	2			+	+		
IAEA_RC-conf	2			+			+
IAF_RC-conf	2		+	+			
IAV_RC-conf	2	+		+			
IAV_RC-elem	2	o		o			
VB_RC-conf	2	+				+	
VD_RC-conf	2	+			+		
VEA_RC-conf	2	+					+
VF_RC-conf	2	+	+				
BDF_RC-conf	3		+		+	+	
EABD_RC-conf	3				+	+	+
EABD_RC-elem	3				o	o	o
IAVF_RC-conf	3	+	+	+			
IAVF_RC-elem	3	o	o	o			
EABDF_RC-conf	4		+		+	+	+
EABDF_RC-elem	4		o		o	o	o
IAVFD_RC-conf	4	+	+	+	+		
IAVFD_RC-elem	4	o	o	o	o		
EABDFIA_RC-conf	5		+	+	+	+	+
EABDFIA_RC-elem	5		o	o	o	o	o
IAVFDB_RC-conf	5	+	+	+	+	+	
IAVFDB_RC-elem	5	o	o	o	o	o	
RC-conf	6	+	+	+	+	+	+
RC-conf_RC-conf	6	+	+	+	+	+	+
RC-conf2_RC-conf	6	+	+	+	+	+	+
RC-elem	6	o	o	o	o	o	o
RC-elem_RC-conf	6	+	+	+	+	+	+
RC-elem_RC-elem	6	o	o	o	o	o	o

Table S2: The 63 pair comparisons to evaluate the similarity between the odor of the reference RC-conf or RC-elem mixture and the odor of a sample including either a single odorant, a sub-mixture or the RC-conf or RC-elem mixtures (controls).

Pair comparison	Stimulus1	Stimulus2	Similarity on a 0-10 scale
1	B_RC-conf	RC-conf	1.71
2	D_RC-conf	RC-conf	1.12
3	EA_RC-conf	RC-conf	0.69
4	F_RC-conf	RC-conf	1.57
5	IA_RC-conf	RC-conf	3.21
6	V_RC-conf	RC-conf	2.37
7	B_RC-conf	RC-elem	4.72
8	D_RC-conf	RC-elem	4.65
9	EA_RC-conf	RC-elem	1.97
10	F_RC-conf	RC-elem	2.97
11	IA_RC-conf	RC-elem	2.67
12	V_RC-conf	RC-elem	2.8
13	B_RC-elem	RC-elem	4.78
14	D_RC-elem	RC-elem	4.75
15	EA_RC-elem	RC-elem	1.18
16	F_RC-elem	RC-elem	2.8
17	IA_RC-elem	RC-elem	1.97
18	V_RC-elem	RC-elem	1.29
19	BEA_RC-conf	RC-conf	2.45
20	DB_RC-conf	RC-conf	2.66
21	DEA_RC-conf	RC-conf	2.28
22	FB_RC-conf	RC-conf	2.86
23	FD_RC-conf	RC-conf	2.53
24	FEA_RC-conf	RC-conf	1.77
25	IAB_RC-conf	RC-conf	4.3
26	IAD_RC-conf	RC-conf	5.68
27	IAEA_RC-conf	RC-conf	4.59
28	IAF_RC-conf	RC-conf	5.77
29	IAV_RC-conf	RC-conf	6.61
30	VB_RC-conf	RC-conf	3.31
31	VD_RC-conf	RC-conf	3.86
32	VEA_RC-conf	RC-conf	3.41

Pair comparison	Stimulus1	Stimulus2	Similarity on a 0-10 scale
33	VF_RC-conf	RC-conf	3.85
34	EAB_RC-conf	RC-conf	2.14
35	IAV_RC-conf	RC-conf	6.07
36	EAB_RC-conf	RC-elem	4.89
37	IAV_RC-conf	RC-elem	3.13
38	EAB_RC-elem	RC-elem	5.65
39	IAV_RC-elem	RC-elem	2.9
40	BDF_RC-conf	RC-conf	4.11
41	EABD_RC-conf	RC-conf	2.61
42	IAVF_RC-conf	RC-conf	6.6
43	EABD_RC-conf	RC-elem	6.78
44	IAVF_RC-conf	RC-elem	3.43
45	EABD_RC-elem	RC-elem	7.04
46	IAVF_RC-elem	RC-elem	3.48
47	EABDF_RC-conf	RC-conf	4.54
48	IAVFD_RC-conf	RC-conf	7.28
49	EABDF_RC-conf	RC-elem	5.21
50	IAVFD_RC-conf	RC-elem	3.06
51	EABDF_RC-elem	RC-elem	6.68
52	IAVFD_RC-elem	RC-elem	6.57
53	EABDFIA_RC-conf	RC-conf	5.76
54	IAVFDB_RC-conf	RC-conf	7.41
55	EABDFIA_RC-conf	RC-elem	4.29
56	IAVFDB_RC-conf	RC-elem	2.76
57	EABDFIA_RC-elem	RC-elem	7.53
58	IAVFDB_RC-elem	RC-elem	7.22
59	RC-conf_RC-conf	RC-conf	7.68
60	RC-elem_RC-conf	RC-conf	3.66
61	RC-conf_RC-conf	RC-conf	7.69
62	RC-elem_RC-conf	RC-elem	7
63	RC-elem_RC-elem	RC-elem	7.26

Table S3: Normalized 21 structural parameters of the 6 odorants obtained with the Dragon software.

Odorant	nCIR	ZM1	GNar	S1K	piPC08	MATS1v	MATS7v	GATS1v	Eig05_EA(ed)	SM02_EA(ed)	SM03_EA(dm)
V	1	0.63	0.97	0.41	0.70	0.89	1	0	0.79	0.70	0
F	1	0.71	1	0.59	1	1	0.45	0.14	0.91	0.65	0.91
IA	0	0.29	0.27	0.54	0	0.36	0.01	0.99	0.32	0.31	1
D	1	1	0.80	0.95	0.95	0.88	0.18	0.39	0.95	1	0
B	1	1	0.80	1	0.85	0.88	0	0.52	1	0.97	0
EA	0	0	0	0	0	0	0.49	1	0	0	1

Odorant	SM10_EA(dm)	SM13_EA(dm)	SpMin3_Bh(v)	RDF035v	G1m	G1v	G1e	G3s	R8u+	nRCOSR
V	0.04	0	0.54	0.30	0.42	0.59	0.61	0.51	0	0
F	0.87	0.87	0.69	0.18	1	0.29	0.34	0	1	0
IA	1	1	0.72	0.23	0.27	0.34	0.39	0.32	0.89	0
D	0	0	0.96	1	0.56	0	0.27	1	0.91	0
B	0.36	0	1	0.97	0	0.68	0	0.10	0.71	0
EA	1	1	0	0	0.69	1	1	0.33	0	0

III. Additional results and partial discussion

The angle distance model well predicted perceptual similarity ($r = 0.63$, $p < 0.001$) in our benchmark dataset that includes mixtures with different levels of complexity, using component odors at various intensities, and eliciting both elemental and configural percepts. The ratio-weighted angle distance model succeeded better at the predictive task ($r = 0.75$, $p < 0.001$). Thus, taking into account components' ratio along with their structural parameters improved not only the similarity prediction performances but also allowed to account for the elemental or configural processing of complex odor mixtures by the olfactory system.

We further applied the ratio-weighted angle distance model on the chemical and sensory data of the 16 wines dataset. To do so, we used the odorants composition and sensory profiles of the wines from Villière et al. (**Data paper included in this manuscript**), and we followed the methodology described in Roche et al. (**Article 1 of this manuscript**) to collect the structural parameters of the molecules from the wines using the Dragon software. We applied the ratio-weighted angle distance model while taking into account the proportion of the molecules in the wines. To be closer to perceptual aspects of this ratio, the proportions were based of the NIF scores of the odorants obtained from the GC-O analysis of the wines. NIF scores corresponded to the proportion of panelists who perceived the odorant. The last step was to translate the sensory profiles of the wines into similarity scores. From the sensory profile of the wines, we inferred similarity values using the Ruzicka similarity. The similarity calculation was carried out following the Equation 3.1.

Equation 3.1:

$$\text{Similarity} = \frac{\sum \min\{x_i, y_i\}}{\sum \max\{x_i, y_i\}}$$

For example, if we consider two wines evaluated on 15 odor sensory attributes (OSA) as shown in Table 3.1, the similarity between the two wines will be equal to $\frac{\sum \min\{OSA_1,\ldots,OSA_{20}\}}{\sum \max\{OSA_1,\ldots,OSA_{20}\}} = \frac{0.25+0.19+1.29+1.02+0.80+0.78+0.26+0.64+0.94+0.62+0.19+0.51+0.55+0.13+1.61}{0.45+0.46+1.51+1.20+0.82+1.20+0.38+0.73+0.95+0.68+0.33+0.61+0.94+0.15+2.23}$, which resulted in a similarity value equal to 0.77.

Table 3.1: Example of the sensory evaluation of two wines on 15 odor sensory attributes.

The two wines are from the 16 wines dataset. Sensory scores varying between 0 and 10.

Wine	Bell pepper	Blackcurrant bud	Blackcurrant fresh	Cherry cooked	Cherry fresh	Cherry stone	Cut grass	Leather	Prune	Smoky	Strawberry fresh	Toasty	Vanilla	Violet	Woody
PN1	0.45	0.46	1.51	1.02	0.80	0.78	0.26	0.64	0.95	0.62	0.33	0.61	0.94	0.15	2.23
PN2	0.25	0.19	1.29	1.20	0.82	1.20	0.38	0.73	0.94	0.68	0.19	0.51	0.55	0.13	1.61

The results showed a high correlation between the estimated similarity (Ruzicka) and the angle distances obtained from the ratio-weighted model ($r = -0.72$, $p < 0.001$, Figure 3.1 black line). Nevertheless, on the whole dataset, the results were mostly driven by the same-wine comparisons (dots at the (0;1) coordinates). When excluding this bias, the correlation remained significant but the predicting abilities drastically dropped ($r = -0.25$, $p < 0.005$, Figure 3.1 red line).

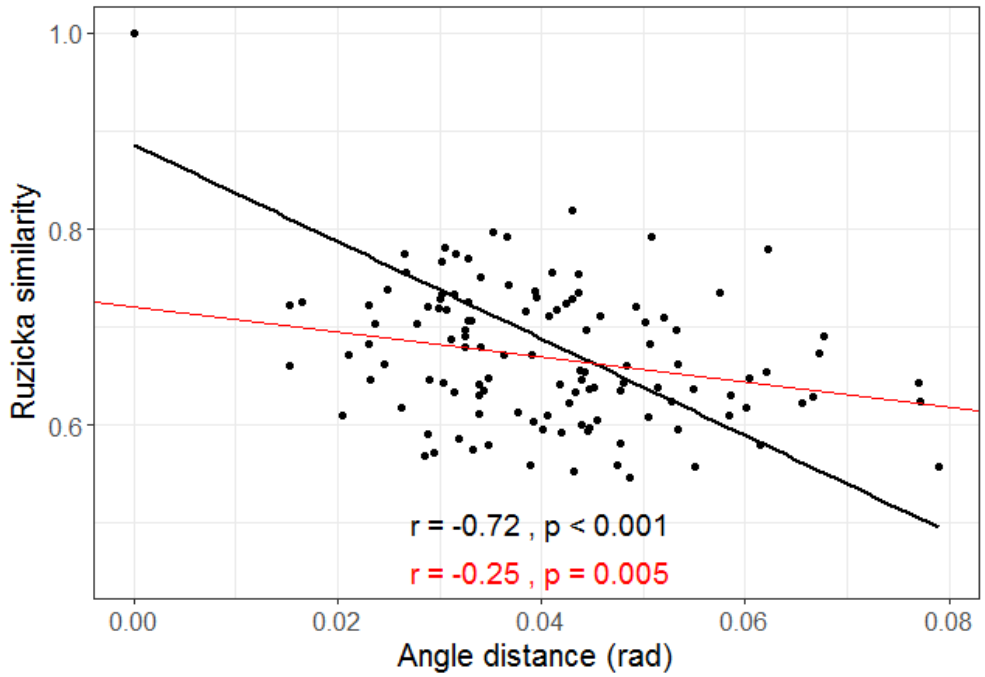


Figure 3.1: Ratio-weighted angle distance model applied to the odorants' composition of wines. Molecular structure vectors were weighted by the NIF scores of the odorants within mixtures. Each dot reflects a comparison between two wines. A linear regression was performed on the entire dataset (black line) and another one without the same-wine comparisons, i.e. dots at the (0;1) coordinates (red line).

However, we should keep in mind some limitations which may account for the observed results. Firstly, experimental sensorysimilarity was not assessed between the wines. Indeed similarity data were inferred from the odor sensory profiles, which may explain why the similarity scores range from 0.6 to 0.8, namely a low variation range. Secondly, in contrast to the study with the 6-component mixtures (Roche et al., **Article 1 of this manuscript**), for the wines we used the NIF scores of odorants as we did not have the intensity or the ratio of each

odorant available. Subsequent work might focus on coupling structural approaches, such as the angle distance model and its extended versions, with intensity predictive models (e.g. Chastrette et al., 1998) to be able to integrate in a more accurate manner the intensity of the components of complex mixtures.

Nevertheless, the results obtained in this first study have shed light on the possibility to efficiently predict the perceptual outcome of complex odorants' mixtures on the basis of their molecular structure. This work focused on predicting the similarity among mixtures of odorants as a first step to predict perceptual features of complex odor mixtures. Still, the aim of this thesis was to predict what a mixture of odorants will smell in verbal terms with quantitative data as in an odor profile. Thus, the two following chapters are related to innovative strategies to predict how odorants' mixtures will smell by integrating flavorists' expertise in the modelling framework.

SECOND AXIS:

EXPERTISE INTEGRATION APPROACH

CHAPTER 4: ONTOLOGY FOR THE ODOR PERCEPTUAL SPACE

I. Introduction

Odorants are commonly characterized using three dimensions: intensity, pleasantness, and odor quality, namely the type of percept. Whereas intensity and pleasantness can be rated using various numeric scales, the odor percept is qualified by semantic labels and thus can be described by a more or less large number of descriptors such as “apple”, “green”, “jasmin”, “meat”, “woody”… Several studies focused on odor descriptors (OD) to highlight the categorical dimensions of the Human odor space, for instance by identifying correlations and relationships among OD in databases gathering descriptions of odorants (Zarzo & Stanton, 2006; Castro et al., 2013; Kumar et al., 2015).

However, the OD dimension alone is not representative of the entire olfactory space, especially because most natural odors are caused by complex mixtures of odorants. The non-linear integration mechanisms that underpin the perception of such mixture have prevented the investigation of the link between chemical composition and odor quality perception. In this chapter, we proposed a knowledge-representation approach based on an ontology to link the perceptual outcome of complex mixtures of odorants to the odor of the single odor compounds, namely odorants. The ontology was developed around three odor description corpuses: odor descriptors (OD) used for single odorants, odor sensory attributes (OSA) used to qualify complex odor objects for instance in odor profiling techniques, and odor qualities (OQ) which is the corpus of specific terms used by experts of odor creation. Indeed, we observed that flavorists did not use OD or OSA when creating a specific flavor but they relied on combinations of OQ such as “fruity”, “floral”, “smoky”...

The construction of the ontology for the odor perceptual space (OOPS) is detailed in the dedicated **ARTICLE 2**. The expertise embedded in the ontology modelling was further used to predict the OSA profile of wines on the basis of their odorants’ composition.

II. ARTICLE 2

OOPS, the Ontology for Odor Perceptual Space: from molecular composition to sensory attributes of odor objects

Alice Roche[1], Nathalie Perrot[2], Thierry Thomas-Danguin[1]

1. CSGA, UMR1324 INRA, UMR6265 CNRS, AgroSup Dijon, Université de Bourgogne Franche - Comté, Dijon, France
2. GMPA, UMR782 INRA, AgroParisTech, Université Paris Saclay, Thiverval-Grignon, France

* Correspondence to be sent to: thierry.thomas-danguin@inra.fr

This manuscript is in preparation.

Keywords: odor perceptual space, odor quality, odor descriptor, odorant, wine, expert knowledge

Abstract

When creating a flavor to elicit a specific odor object characterized by odor sensory attributes (OSA), expert perfumers or flavorists use mental combinations of odor qualities (OQ) such as Fruity, Green, Smoky. However, OSA and OQ are not directly related to the molecular composition in terms of odorants that constitute the chemical stimuli supporting odor object perception because of the complex non-linear integration of odor mixtures within the olfactory system. Indeed, single odorants are described with odor descriptors (OD), which can be found in various databases. Although classifications and aroma wheels studied the relationships between OD and OQ, the results are highly dependent of the studied products. Nevertheless, ontologies have proved to be very useful in sharing concepts across applications in a generic way but also to allow experts' knowledge integration implying non-linear cognitive processes. In this paper we constructed the Ontology for Odor Perceptual Space (OOPS) to merge OD into a set of OQ best characterizing the odor further translated in a set of OSA thanks to expert knowledge integration. Results showed that OOPS can help to bridge molecular composition to odor perception and description as demonstrated in the case of wines.

Introduction

Within the physical world, colors are characterized by light wavelength, tones by sound frequency, and odors by the chemical composition of the stimulus. Within the perceptual space, colors are defined by specific words like red or blue, tones are referred to by dedicated notes like C or E ♭ , while odors are usually identified by naming their sources like rose or lemon (Dubois & Rouby, 2002). Therefore, if colors and tones can be well defined experimentally, odors are difficult to describe with a consensual vocabulary but also difficult to measure physically because they mostly results from the coding, by the olfactory system, of complex mixtures of odorants, which are volatile organic compounds varying in chemical nature and concentration (Thomas-Danguin et al., 2014).

Olfactory coding of complex mixtures of odorants induces perceptual interactions, which can take place at several steps of the olfactory information processing, and the overall is not a simple sum of the odors of each odorant embedded in the mixture (Thomas-Danguin et al., 2014). Synergy and masking effects have been often reported (Cain & Drexler, 1974; Ishii et

al., 2008; Ferreira, 2012a; Tempere et al., 2016), but also perceptual dominance (Ferreira, 2012b), or configural and elemental perception (Jinks & Laing, 2001; Howard & Gottfried, 2014; Romagny et al., 2018). For instance, a ternary mixture, composed of three odorants respectively described as "strawberry", "caramel" and "violet", elicits, at a specific proportion of each compound, the perception of a "pineapple" odor (Le Berre et al., 2008). The mechanisms behind these perceptual interactions are not well understood and still poorly investigated. Consequently, the description of the perceptual outcome of a complex mixture using odor sensory attributes (OSA) is not straightforward, but sensory descriptive analysis using OSA has been shown to be one of the most powerful, sophisticated and most extensively used tools in sensory science to provide a complete description of the sensory characteristics of a product (Varela & Ares, 2012). The global odor percept is especially hardly predictable on the basis of the mixtures' chemical composition, namely each single odorant own odor, described with odor descriptors (OD).

Several databases are compiling OD of large sets of odorants: Arctander's handbook (Arctander, 1969), Atlas of odor character profiles (Dravnieks, 1985), Fenaroli's handbook (Burdock, 2010), Flavor-Base (Leffingwell & Associates, http://www.leffingwell.com/flavbase.htm), Flavornet (Arn & Acree, 2004), Flavors and Fragrances of Sigma-Aldrich (http://www.sigmaaldrich.com/industries/flavors-and-fragrances.html), The good scents company (Luebke, 1980). If the use of semantic descriptors was found to be a stable method to provide the most reproducible characterization of odorous substances (Dravnieks, 1982), the vocabulary used in describing odor is extensive and ambiguous. As a matter of fact, are "citrus odor" and "odor of citrus" referring to the same odor descriptors? (Barkat-Defradas & Motte-Florac, 2016). Moreover, there is no agreement about the number of OD essential to cover the complete range of odor stimuli which varies from 4 to 146 (Chastrette, 2002). Though several teams worked on the different relationships, associations, or similarities between OD, none of them had yet gained wide acceptance (Chastrette 1988; Zarzo and Stanton, 2006; Kumar et al., 2015).

In most cases, it is not possible to make a direct link between the OD of the odorants released from an odor source, e.g. a food product, and its perceived odor. This is probably the reason why experts, for instance flavorists who create specific odors from combinations of odor active raw materials such as molecules, are not using odorants' OD but a rather different set of descriptors to organize their practical knowledge acquired along with experience (Langlois et al., 2011). Indeed, to conceptualize the perception of specific odor traits of an odor source,

flavorists use a set of odor qualities (OQ) that they can mentally combine, probably through mental imagery (Tempere et al., 2014). For example, according to an expert flavorist, the OSA "Cherry cooked" is composed of the OQ "Almond", "Cooked", "Floral", "Fruity", "Green", "Peel" and "Spicy". The OQ may be considered as "blocks", where each block could be composed of several odorants referring to different OD (e.g. Jaubert et al., 1987). In a sense, OQ refer to broad odor categories, related more to odor material than to molecules (Zarzo & Stanton, 2009). Classifications and flavor wheels usually dedicated to a specific category of food products such as wine have been established and could help to make links between OD and OQ. However, these classifications are highly dependent of the studied databases and/or food product and are hardly reconcilable (Dravnieks, 1985; Zarzo and Stanton, 2006; caramel: Paravisini et al., 2014; honey: International Honey Commission (IHC) http://www.ihc-platform.net/reports.html; wine: Noble et al., 1987). For example, whereas the OD "Apple" is classified in the OQ "Fruity" in the five above cited sources, the OD "Vanilla" is classified in five different OQ "Spicy", "Balsamic", "Warm", "Wood/Phenolic" or "Caramel/Vanilla" depending on the source.

To overcome these issues, this paper had for aim to use the ontology approach to make the link between OSA, the odor sensory attributes used to describe a given odor source; OQ, the concepts manipulated by experts to conceptualize odors; and OD, the odor descriptors used to qualify odorants. With the help of expert flavorists, we developed and formalized the Ontology for Odor Perceptual Space (OOPS) to organize the vocabulary of the odor perceptual space and to describe the relationships between the OD, OQ and OQ. The aim was to merge the information expressed by OD in order to formally characterize odors into a conceptual and generic annotation of OQ, namely not associated to a specific food product. We further used the OOPS to predict the odor profiles of two red wines, that is to say the OSA used by a trained panel to describe these wines (Villière et al., **Data paper included in this manuscript**).

Materials and methods

Wines

Villière et al. (**Data paper included in this manuscript**) studied the sensory profiles and the chemical composition in terms of odor-active compounds of 16 red wines, varying according to their exemplarity for the grape variety (Loison et al., 2015). The experimental factors of the wines are listed in Table 1. Sensory profiles resulted in the identification of 15 discriminant OSA between the wines according to their grape varieties (Table 2). The results of Gas Chromatograpy - Mass Spectrometry - Olfactometry (GC-MS-O) analyses led to identify 46 odorant zones (molecules and mixtures of molecules) which corresponded to 49 identified odorants (Table 3). Raw data are available on an open-source repository (Villière et al., 2018).

Table 1: Wines experimental factors. List of the 16 red wines and their experimental factors: grape variety, vintage, PDO (Protected Designation of Origin) and exemplarity of the wine to its grape variety rated on a 10 point scale.

Wine	Grape variety	Vintage	PDO	Exemplarity
PN-A	Pinot Noir	2009	Ladoix	6.26
PN-B	Pinot Noir	2009	Côte de Nuits-Villages	6.21
PN-C	Pinot Noir	2010	Maranges	6.01
PN-D	Pinot Noir	2010	Bourgogne	5.78
PN-E	Pinot Noir	2009	Bourgogne	4.54
PN-F	Pinot Noir	2009	Bourgogne Hautes Côtes de Beaune	4.48
PN-G	Pinot Noir	2009	Savigny-lès-Beaune	4.38
PN-H	Pinot Noir	2009	Bourgogne	4.19
CF-A	Cabernet Franc	2010	Bourgueil	7.12
CF-B	Cabernet Franc	2010	Bourgueil	6.54
CF-C	Cabernet Franc	2009	Chinon	6.49
CF-D	Cabernet Franc	2010	St-Nicolas-de-Bourgueil	6.35
CF-E	Cabernet Franc	2010	Bourgueil	6.25
CF-F	Cabernet Franc	2010	Saumur	4.09
CF-G	Cabernet Franc	2010	Bourgueil	3.84
CF-H	Cabernet Franc	2010	Chinon	3.26

Table 2: List of the 15 odor sensory attributes (OSA) discriminant between the 16 wines according to their grape varieties.

Bell pepper	Prune
Blackcurrant bud	Smoky
Blackcurrant fresh	Strawberry fresh
Cherry cooked	Toasty
Cherry fresh	Vanilla
Cherry stone	Violet
Cut-grass	Woody
Leather	

Table 3: Molecular space of the 16 red wines identified by GC-MS-O. List of the 49 odorants identified across the 16 wines and identified by their CAS number and name.

CAS	Odorant
4312-99-6	1-Octen-3-one
431-3-8	2,3-Butanedione
600-14-6	2,3-Pentanedione
91-10-1	2,6-Dimethoxyphenol
90-05-1	2-Methoxyphenol
110-19-0	2-Methylpropyl acetate
620-17-7	3-Ethylphenol
24683-00-9	3-Isobutyl-2-methoxypyrazine
25773-40-4	3-Isopropyl-2-methoxypyrazine
51755-83-0	3-Mercapto-1-hexanol
123-51-3	3-Methyl-1-butanol
590-86-3	3-Methylbutanal
123-92-2	3-Methylbutyl acetate
2785-89-9	4-Ethyl guaïacol
123-07-9	4-Ethylphenol
626-89-1	4-Methyl-1-pentanol
75-07-0	Acetaldehyde
64-19-7	Acetic acid
100-52-7	Benzaldehyde
122-78-1	Benzene acetaldehyde
60-12-8	Benzene ethanol
100-51-6	Benzene methanol
107-92-6	Butyric acid
96-48-0	Butyrolactone
334-48-5	Decanoic acid

CAS	Odorant
75-18-3	Dimethyl sulfide
64-17-5	Ethanol
141-78-6	Ethyl acetate
105-54-4	Ethyl butanoate
110-38-3	Ethyl decanoate
106-33-2	Ethyl dodecanoate
123-66-0	Ethyl hexanoate
106-32-1	Ethyl octanoate
105-37-3	Ethyl propanoate
7452-79-1	Ethyl-2-methylbutanoate
97-62-1	Ethyl-2-methylpropanoate
108-64-5	Ethyl-3-methylbutanoate
142-62-1	Hexanoic acid
503-74-2	Isovaleric acid
108-39-4	m-Cresol
74-93-1	Methanethiol
3268-49-3	Methional
505-10-2	Methionol
80-62-6	Methyl-2-methylpropenoate
106-44-5	p-Cresol
103-45-7	Phenethyl acetate
108-95-2	Phenol
7446-09-5	Sulphur dioxide
39212-23-2	Whyskeylactone

Elicitation of odor qualities (OQ) by expert flavorists

Four senior flavorists participated in the expert knowledge collection. The elicitation process was based on a 1-hour private guided interview. Flavorists were not aware of the studied food matrix in order to collect unbiased data regarding the food product.

The experts received monadically the 15 OSA used in the wines' sensory profiles (Table 1) and were asked i) if the OSA was composed of a single OQ or of more than one OQ and ii) in case the considered OSA was composed of several OQ, to enumerate the OQ that were needed to construct the OSA. We aggregated the information of the four flavorists following Equation 1, OSA being a given odor sensory attribute, Exp1[OQ(OSA)], Exp2[OQ(OSA)], Exp3[OQ(OSA)] and Exp4[OQ(OSA)] being the sets of OQ used to describe an OSA by the four experts.

Equation 1: $OSA = Exp1[OQ(OSA)] \cup Exp2[OQ(OSA)] \cup Exp3[OQ(OSA)] \cup Exp4[OQ(OSA)]$

As a result, we obtained a binary matrix made of in rows the 20 OQ elicited (Almond, Cooked, Cut-Grass, Floral, Fresh, Fruity, Green, Honey, Lactonic, Leather, Peel, Smoky, Spicy, Sulfurous, Toasty, Vanilla, Vegetable, Violet, Wine-like and Woody) and in columns the target OSA (Table 4).

Table 4: Link between the 20 OQ (rows) and the 15 OSA (columns), represented as a binary matrix. The value 1 indicates that the OQ was part of the composition of the OSA according to the experts.

OQ	Bell pepper	Blackcurrant bud	Blackcurrant fresh	Cherry cooked	Cherry fresh	Cherry stone	Cut-grass	Leather	Prune	Smoky	Strawberry fresh	Toasty	Vanilla	Violet	Woody
Almond				**1**	**1**	**1**			**1**						
Cooked				**1**	**1**	**1**			**1**		**1**				
Cut-grass							**1**								
Floral	**1**	**1**	**1**	**1**	**1**	**1**					**1**				
Fresh	**1**	**1**	**1**												
Fruity		**1**	**1**						**1**		**1**				
Green	**1**	**1**	**1**	**1**	**1**	**1**					**1**				
Honey									**1**						
Lactonic									**1**						
Leather								**1**							
Peel				**1**	**1**	**1**									
Smoky										**1**					
Spicy				**1**	**1**	**1**									
Sulfurous	**1**	**1**	**1**												
Toasty	**1**											**1**			
Vanilla		**1**											**1**		
Vegetable	**1**														
Violet															
Wine-like		**1**	**1**												
Woody															**1**

Quantitative description of the odorants

We compiled the data of three databases to collect the odor descriptors (OD) of the 49 odorants identified in the wines: Arctander's handbook (3102 chemicals described by Steffen Arctander himself), Flavor-Base (commercially available Leffingwell & Associates database, marketed as Flavor-Base Pro © 2010, flavor descriptions collected from many sources over the course of more than 40 years) and The good scents company (publicly available database, the odor descriptions from one to several sources are listed in the "Organoleptic Properties" section).

We extracted manually the OD from these databases. The words describing the odorants were tokenized. Suffixes (e.g. "like", "note"), auxiliary verbs (e.g. "has") and some other words

that did not refer to olfactory information (e.g. “powerful”) were discarded. Unlike the analysis of the Arctander database proposed by Chastrette et al. (1988), we kept all the OD into account and we did not combine very similar descriptors (like Leather/Leathery or Wine/Winey) For instance, the odor of Ethyl butanoate (CAS 105-54-4) was specified in Arctander as “Powerful, ethereal-fruity odor suggestive of Banana and Pineapple, and very diffusive” these annotations resulted in the set of OD: “ethereal-fruity”, “banana” and “pineapple”.

Then we created the OD database by aggregating the information of the three databases following Equation 2, M being a given odorant, Arct[OD(M)], Flavor-Base[OD(M)] and Goodscent[OD(M)] being the sets of OD of the odorant M by the Arctander, Flavor-Base and Goodscent databases. We ended up with 175 different OD for the 49 odorants.

Equation 2: $OD\ database(M) = Arct[OD(M)] \cup FlavorBase[OD(M)] \cup Goodscent[OD(M)]$

For a given odorant, the description was thus provided by the OD database as a set of terms in which each item may be associated to an “intensity”. We defined this intensity as the number of citation of the same OD for a given odorant across the databases: the higher the number of citation, the more “intense” the smell related to this OD was expected for the odorant. As an example, the odorant description of Ethyl butanoate was {ethereal-fruity; banana; pineapple} by Arctander, {ethereal; fruity; buttery; pineapple; banana; ripe fruit; juicy} by Flavor-Base and {fruity; juicy; pineapple; cognac} by The good scents company. The resulting quantitative description of Ethyl butanoate in the OD database was the following: OD(Ethyl butanoate) = [(banana, 2); (buttery, 1); (cognac, 1); (ethereal, 1); (ethereal-fruity, 1); (fruity, 2); (juicy, 2); (pineapple, 3); (ripe fruit, 1)].

Relationships between odor descriptors (OD) and odor qualities (OQ)

The correspondence between an OD and one or several OQ was obtained thanks to the expertise of a junior flavorist. This expert was not one of the four flavorists previously interviewed for OQ elicitation. The methodology used to obtain the relationships was based on a check-all-that-apply (CATA) questionnaire (Dooley et al., 2010). The CATA list

consisted of the 20 OQ defined by the experts during the elicitation step (see 2.2 above). For each OD of the OD database, the flavorist was asked if the OD supported none, one, or several OQ. For instance for the OD "Apple", the flavorist was asked to tick all the OQ that correspond to this OD (e.g. "Fruity").

We obtained a binary matrix with the OQ in columns and OD in rows. These results allowed us to translate each OD sets into OQ sets. For example for Ethyl butanoate, described as OD(Ethyl butanoate) = [(banana, 2); (buttery, 1); (cognac, 1); (ethereal, 1); (ethereal-fruity, 1); (fruity, 2); (juicy, 2); (pineapple, 3); (ripe fruit, 1)], the OD banana, ethereal-fruity, fruity, pineapple and ripe fruit were contributing to the OQ Fruity (Table 5).

Equation 3: $$OQ(M) = \sum Intensity(OD(M) \times OQ)$$

Following the Equation 3 given for an odorant M, we could assume that the OQ set of Ethyl butanoate was the following: OQ(Ethyl butanoate) = [(Almond, 0); (Cooked, 0); (Cut-grass, 0); (Floral, 0); (Fresh, 0); (Fruity, 9); (Green, 0); (Honey, 0); (Lactonic, 0); (Leather, 0); (Peel, 0); (Smoky, 0); (Spicy, 0); (Sulfurous, 0); (Toasty, 0); (Vanilla, 0); (Vegetable, 0); (Violet, 0); (Wine-like, 0); (Woody, 0)].

Table 5: Link between the nine OD of Ethyl butanoate (rows) and the 20 OQ (columns), represented as a binary matrix. The intensity of each OD is specified in a supplementary column.

OD	**Intensity**	Almond	Cooked	Cut-grass	Floral	Fresh	**Fruity**	Green	Honey	Lactonic	Leather	Peel	Smoky	Spicy	Sulfurous	Toasty	Vanilla	Vegetable	Violet	Wine-like	Woody
banana	2	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
buttery	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
cognac	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ethereal	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ethereal-fruity	1	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
fruity	2	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
juicy	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
pineapple	3	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ripe fruit	1	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The Ontology for Odor Perceptual Space (OOPS)

We formalized the Ontology for Odor Perceptual Space (OOPS) as a set {C, R, P}, where C corresponded to the three classes OD, OQ and OSA with respectively 175 sub-classes from the databases aggregation, 20 sub-classes from the expertise collection and 15 sub-classes from the sensory evaluation of the wines; R represented the hierarchical relations among the classes by "is-a" relations; and P, as properties, represented the non-hierarchical associative relations between classes as shown in Figure 1.

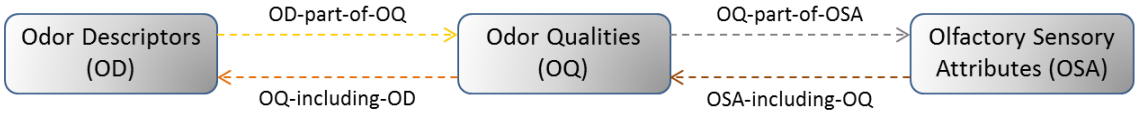


Figure 1: Object properties between the classes OD, OQ and OSA of the OOPS ontology.

Results from the data collection in table forms, were implemented in OWL using the software Protégé (open-source ontology editor, version 5.2.0; Musen, 2015). This allowed the visualization of the properties among the classes OD, OQ and OSA; an example is shown in Figure 2 for the OQ "Vanilla". Such representation highlighted that the OD "vanilla" and "tonka" are part of the OQ "Vanilla". Moreover, the OQ "Vanilla" is part of the OSA "VANILLA" and "BLACKCURRANT BUD". From a practical point of view, these relationships illustrated that an odorant described as "vanilla" or "tonka" was part of the OQ category "Vanilla" and should contribute to the perceptual construction of the odor of Vanilla and Blackcurrant bud, which are OSA.

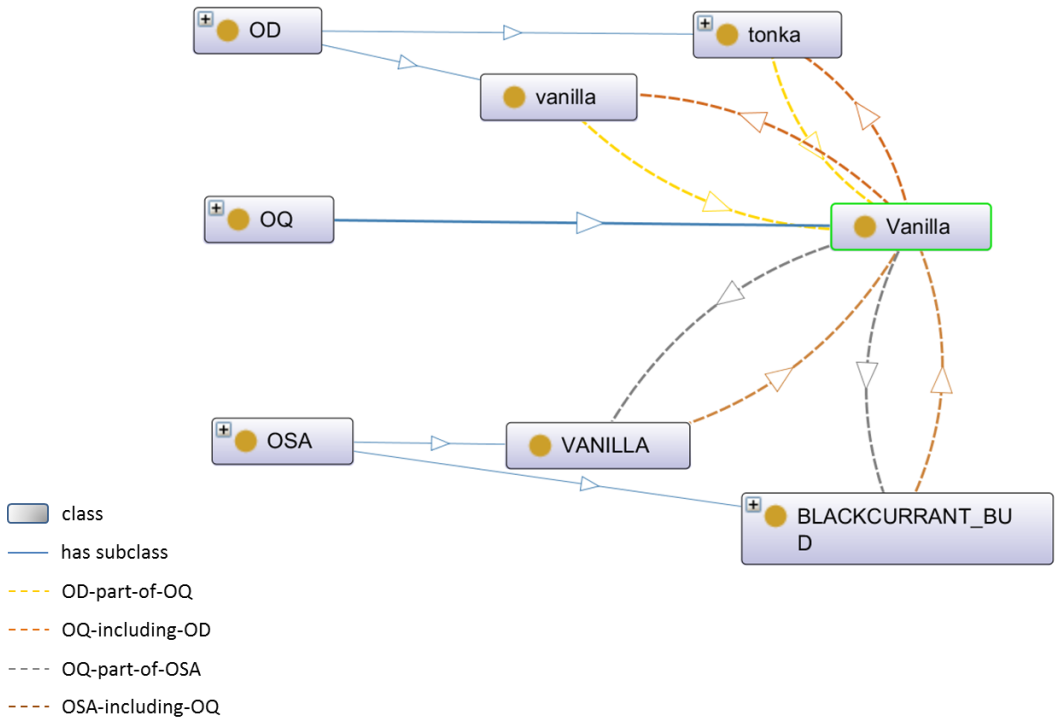


Figure 2: Properties and relationships among the classes OD, OQ and OSA considering the OQ Vanilla.

The implementation of the OOPS in OWL conferred the ability to mine the data through queries such as:

- In which OQ, the OD “almond” is included?

<OQ-including-OD some almond>:"Almond"

- Which OD are parts of the OQ “Almond”?

<OD-part-of-OQ some Almond>:“almond”

- In which OSA, the OQ “Almond” is included?

<OSA-including-OQ some Almond>: “CHERRY_COOKED”, “CHERRY_FRESH”, “CHERRY_STONE”, “PRUNE”

- Which OQ are parts of the OSA “Prune”?

<OQ-part-of-OSA some Prune>: “Almond”, “Cooked”, “Fruity”, “Honey”, “Lactonic”

All together the OOPS led to the fast visualization of relationships among the three classes of odor semantic descriptors OD, OQ and OSA in order to estimate the OQ or OSA profiles of odorants and conversly (Figure 3). As for example with the odorant Ethyl butanoate, described by the OD(Ethyl butanoate) = [(banana, 2); (buttery, 1); (cognac, 1); (ethereal, 1); (ethereal-fruity, 1); (fruity, 2); (juicy, 2); (pineapple, 3); (ripe fruit, 1)], we were able to estimate its contribution to the OQ "Fruity" and then to the OSA "Bell pepper", "Blackcurrant bud", "Blackcurrant fresh", "Cherry cooked", "Cherry fresh", "Cherry stone", "Prune" and "Strawberry fresh".

Intensities related to OD can be propagated along the relationships between OD and OQ as well as between OQ and OSA. The OQ set of Ethyl butanoate was equal to OQ(Ethyl butanoate) = [(Almond, 0); (Cooked, 0); (Cut-grass, 0); (Floral, 0); (Fresh, 0); (Fruity, 9); (Green, 0); (Honey, 0); (Lactonic, 0); (Leather, 0); (Peel, 0); (Smoky, 0); (Spicy, 0); (Sulfurous, 0); (Toasty, 0); (Vanilla, 0); (Vegetable, 0); (Violet, 0); (Wine-like, 0); (Woody, 0)] , as previously mentioned. Regarding the OSA set, we obtained:

OSA(Ethyl butanoate) = [(Bell pepper, 9); (Blackcurrant bud, 9); (Blackcurrant fresh, 9); (Cherry cooked, 9); (Cherry fresh, 9); (Cherry stone, 9); (Cut-grass, 0); (Leather, 0); (Prune, 9); (Smoky, 0); (Strawberry fresh, 9); (Toasty, 0); (Vanilla, 0); (Violet, 0); (Woody, 0)].

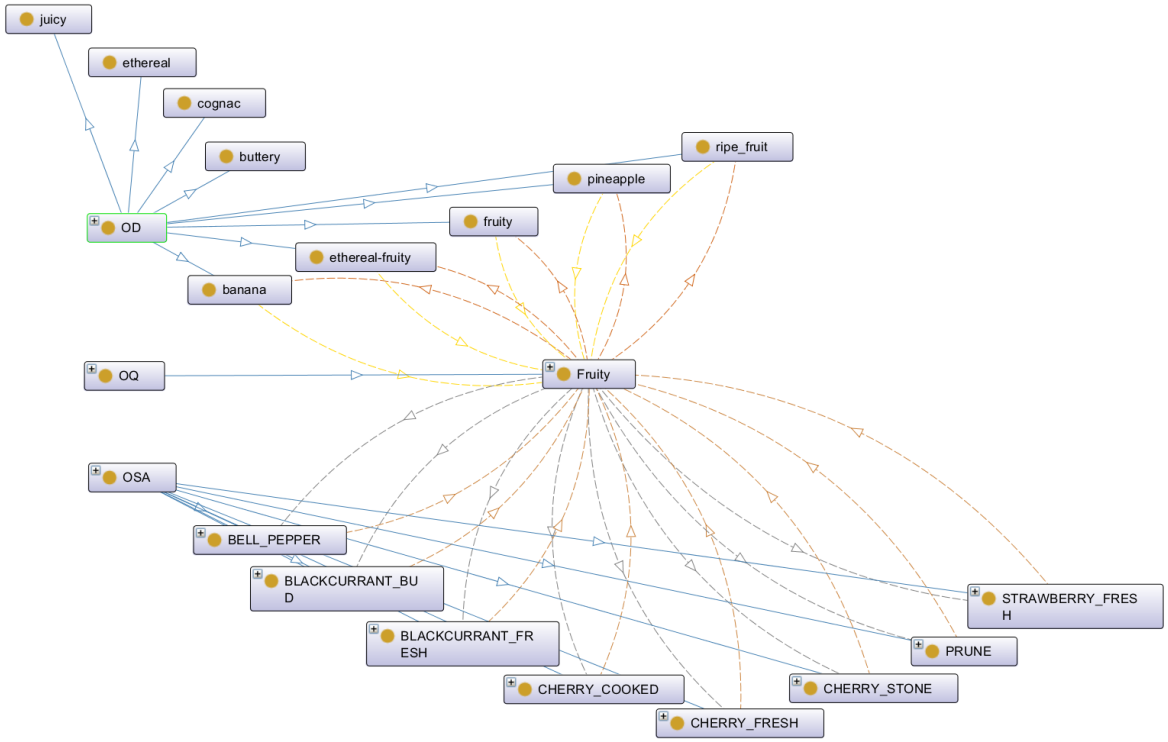


Figure 3: Properties and relationships among the classes OD, OQ and OSA considering the OD of the odorant Ethyl butanoate.

Application of the OOPS to wines

We applied the OOPS to establish the OQ and OSA profiles of two wines starting on the basis of their molecular composition. Two wines were selected among the 16 used by Loison et al. (2015) presented in Table 1: one good example of the grape variety Pinot Noir (PN-A) and one good example of the grape variety Cabernet Franc (CF-A).

We estimated the OQ and OSA sets of each odorant present in these two wines. For a given wine, we summed the OQ and OSA sets of the odorants included in the wine weighted by their intensities.

Firstly, we obtained the OQ profiles of the wines PN-A and CF-A, respectively OQ(PN-A) and OQ(CF-A):

OQ(PN-A) = [(Almond, 1); (Cooked, 3); (Cut-grass, 2); (**Floral, 25**); (Fresh, 1); (**Fruity, 118**); (**Green, 12**); (Honey, 6); (Lactonic, 1); (Leather, 1); (Peel, 4); (**Smoky, 24**); (Spicy, 10);

(Sulfurous, 3); (Toasty, 2); (Vanilla, 4); (Vegetable, 8); (Violet, 0); (Wine-like, 9); (Woody, 5)]

OQ(CF-A) = [(Almond, 3); (Cooked, 4); (Cut-grass, 1); (**Floral, 20**); (Fresh, 1); (**Fruity, 97**); (**Green, 15**); (Honey, 3); (Lactonic, 0); (Leather, 4); (Peel, 4); (**Smoky, 20**); (Spicy, 1); (Sulfurous, 4); (Toasty, 0); (Vanilla, 0); (**Vegetable, 21**); (Violet, 0); (Wine-like, 10); (Woody, 4)]

Values in bold corresponded to OQ with an intensity higher than 5% of the total intensity of the OQ in the corresponding wine. At this step, the two wines were described as Fruity wines with Floral, Green and Smoky notes, and CF-A mainly differed from PN-A according to its Vegetable note.

We estimated the OSA profiles of the two wines PN-A and CF-A, respectively OSA(PN-A) and OSA(CF-A):

OSA(PN-A) = [(**Bell pepper, 51**); (**Blackcurrant bud, 172**); (**Blackcurrant fresh, 168**); (**Cherry cooked, 55**); (**Cherry fresh, 55**); (**Cherry stone, 55**); (Cut-grass, 2); (Leather, 1); (**Prune, 129**); (Smoky, 24); (**Strawberry fresh, 158**); (Toasty, 2); (Vanilla, 4); (Violet, 0); (Woody, 5)]

OSA(CF-A) = [(**Bell pepper, 61**); (**Blackcurrant bud, 147**); (**Blackcurrant fresh, 147**); (**Cherry cooked, 47**); (**Cherry fresh, 47**); (**Cherry stone, 47**); (Cut-grass, 1); (Leather, 4); (**Prune, 107**); (Smoky, 20); (**Strawberry fresh, 136**); (Toasty, 0); (Vanilla, 0); (Violet, 0); (Woody, 4)]

Values in bold corresponded to OSA with an intensity higher than 5% of the total intensity of the OSA in the corresponding wine. From these OSA sets we were able to point out differences among the two wines (Figure 4). The PN-A wine was predicted to have higher intensity of the OSA Cut-grass, Toasty and Vanilla and lower intensity of the OSA Bell pepper and Leather than the CF-A wine. These results were consistent with the literature because PN and CF wines are described as Fruity wines, while CF wines are usually described as having a Bell pepper specific odor note (Lawrence et al., 2013).

According to the actual sensory profiles of the wines obtained with a trained panel (Villière et al., **Data paper included in this manuscript**), PN-A was perceived as more Toasty and Vanilla than CF-A which was also predicted using the OOPS approach. However some

discrepancies between the wines odor profile could be noted. For instance, the sensory evaluation the CF-A wine revealed a higher intensity of the OSA Cut-grass and a lower intensity of the OSA Leather compared to the PN-A wine, while the predicted values obtained with the OOPS approach suggested the opposite situation.

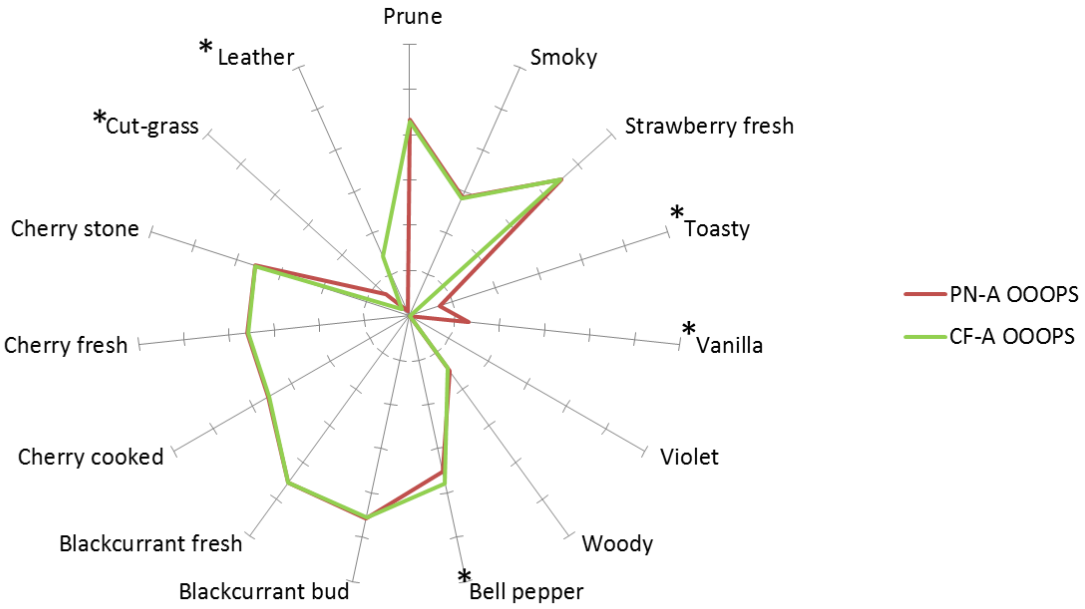


Figure 4: Radar plots illustrating the profile of the OSA in the PN-A and CF-A wines. Proportions of OSA are plotted in log scales (arbitrary units). OSA marked with * indicate significantly different intensities between the two wines (5%).

Conclusions and future work

In this paper, we presented the development of the OOPS, the Ontology for Odor Perceptual Space, designed for fixing the vocabulary of the odor perceptual space and the relationships between the different semantic framework involved: OD, OQ and OSA.

An example of application of the OOPS on a food product was presented with the odorant composition of two red wines (Villière et al., **Data paper included in this manuscript**) to estimate their OQ and OSA profiles.

The originality of the OOPS is that it has been developed on the basis of experts' knowledge with the aim to be disconnected from a given odor source of food product. Therefore, it was needed to integrate the specific vocabulary used by experts, that is the Odor Qualities (OQ) semantic space in top of the Odor Descriptors (OD) used for odorants and the Odor Sensory Attributes (OSA) used by descriptive panels. The knowledge formalized in OOPS is supposed to be quite generic and thus the ontology could provide a standard tool for communication among experts to increase knowledge sharing and can be helpful in training sensory panels for odor profiling (Medjkoune et al., 2016).

The real olfactory world is made of a large variety of odor sources that comprise complex olfactory multi-molecular mixtures. For example, coffee, rose or red wine contain hundreds of different odor-active compounds (Thomas-Danguin et al., 2014). Therefore, a central question in olfaction research is to find a metric to be used to predict how such complex mixtures smell in verbal descriptor terms (Snitz et al., 2013). This was precisely the aim of the OOPS. Indeed, one of the strength of the OOPS approach was to provide a tool to estimate a quantitative odor profile of a given olfactory source knowing its chemical composition, i.e. the list of odorants released by the source. Indeed, from the list of molecules it is possible to predict intensities of OSA, which are the odor sensory descriptors of the global odor of the source. In that sense, the OOPS approach is very innovative since up to now, only the similarity between complex mixtures of odorants can be predicted on the basis of their chemical composition but not an *ab initio* sensory profile (Snitz et al., 2013). An example of application of the OOPS on a food product was presented. The sensory profiles of two wines from two different grape varieties were predict using the OOPS and the intensities of OSA were compared to those obtained with a trained descriptive panel (Loison et al., 2015) to estimate their OSA profiles. Overall, the predicted odor profiles of these real food products were mostly in line with experimentally obtained sensory profiles. Nevertheless, two limitations may also be highlighted.

The first limitation concerns the knowledge currently integrated in the OOPS which is rather limited since it has been obtained after experts' knowledge elicitation using a limited number of OSA. Even if several OSA were very common to several odor sources, especially food (e.g. Vanilla, Woody, Cut-grass), it is obvious that it is insufficient to cover the wide range of odors (Bushdid et al., 2014) even considered through a limited number of discrete clusters (Castro et al., 2013; Dunkel et al., 2014). The second limitation is related to odor profiles prediction. If the proof of concept based on wine odor is promising, it is important to keep in

mind that not all the molecules have been identified in the wines considered (Villière et al., **Data paper included in this manuscript**) and the concentration of odorants and their ratio also have an important role in odor mixtures perception (Thomas-Danguin et al., 2014). These two critical factors are not accounted for in the actual ontology approach and should be further investigated.

Consequently, future work will be to increase the data and knowledge embedded in the ontology to allow more complete and accurate predictions. Moreover, further modelling approaches using OOPS may be developed to increase chemistry-based odor mixture prediction performance in order to take into account not only the chemical nature of odorants but also their intensity or concentration within the mixture. The OOPS is likely the first step towards generating a knowledge-based innovative modelling pathway to allow predicting complex odorant mixtures odor in verbal descriptor terms.

Acknowledgments

This work was supported by grants from the Conseil Régional Bourgogne Franche-Comté. AR was supported by an INRA thesis fellowship. We acknowledge the 4 senior flavorists as well as Baptiste Lissarague, junior flavorist, who participated in the expert knowledge collection and Thibaut Dosne for his help in data collection.

References

Arctander, S. 1969. Perfume and flavor chemicals, Volumes I and II. Montclair, NJ, USA.

Arn, H & Acree, TE. 1998. Flavornet: A database of aroma compounds based on odor potency in natural products. Developments in food science, 40:27. doi:10.1016/S0167-4501(98)80029-0.

Atanasova, B, Thomas-Danguin, T, Chabanet, C, Langlois, D, Nicklaus, S & Etiévant, P. 2005. Perceptual interactions in odour mixtures: Odour quality in binary mixtures of woody and fruity wine odorants. Chemical Senses, 30(3):209-217. doi:10.1093/chemse/bji016.

Barkat-Defradas, M & Motte-Florac, E. 2016. Verbalization of olfactory perception. In: Words for odours: Language skills and cultural insights, eds Barkat-Defradas, M & Motte-Florac, E. Cambridge scholars publishing, Newcastle upon Tyne, UK. 1-22.

Burdock, GA. 2010. Fenaroli's handbook of flavor ingredients: Sixth edition. CRC press, Boca Raton, FL, USA.

Bushdid, C, Magnasco, MO, Vosshall, LB & Keller, A. 2014. Humans can discriminate more than 1 trillion olfactory stimuli. Science, 343(6177):1370-1372. doi:10.1126/science.1249168.

Cain, WS & Drexler, M. 1974. Scope and evaluation of odor counteraction and masking. Annals of the New York academy of sciences, 237:427-439. doi:10.1111/j.1749-6632.1974.tb49876.x.

Castro, JB, Ramanathan, A & Chennubhotla, CS. 2013. Categorical Dimensions of Human odor descriptor space revealed by non-negative matrix factorization. Plos one, 8(9):e73289. doi:10.1371/journal.pone.0073289.

Chastrette, M, Elmouaffek, A & Sauvegrain, P. (1988). A multidimensional statistical study of similarities between 74 notes used in perfumery. Chemical Senses, 13(2), 295–305. doi:10.1093/chemse/13.2.295

Chastrette, M. 2002. Classification of odors and structure-odor relationships. In: Olfaction, taste and cognition eds Rouby, C, Schaal, B, Dubois, D, Gervais, R & Holley, A. Cambridge university press, New York, NY, USA. 100-116.

Dooley, L, Lee, YS & Meullenet, JF. 2010. The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. Food quality and preference, 21(4):394-401. doi:10.1016/j.foodqual.2009.10.002.

Dravnieks, A. 1982. Odor quality: Semantically generated multidimensional profiles are stable. Science, 218(4574):799-801. doi:10.1126/science.7134974.

Dravnieks, A. 1985. Atlas of odor character profiles. ASTM, Philadelphia, PA, USA.

Dubois, D. 2006. Des catégories d'odorants à la sémantique des odeurs : Une approche cognitive de l'olfaction. Terrain, 47:89-106. doi:10.4000/terrain.4263.

Dubois, D & Rouby, C. 2002. Names and categories for odors: The veridical label. In: Olfaction, taste and cognition eds Rouby, C, Schaal, B, Dubois, D, Gervais, R & Holley, A. Cambridge university press, New York, NY, USA. 47-66.

Dunkel, A, Steinhaus, M, Kotthoff, M, Nowak, B, Krautwurst, D, Schieberle, P & Hofmann, T. 2014. Natures chemical signatures in human olfaction: A foodborne perspective for future biotechnology. Angewandte reviews, 53(28):7124-7143. doi:10.1002/anie.201309508.

Ferreira, V. 2012a. Revisiting psychophysical work on the quantitative and qualitative odour properties of simple odour mixtures: A flavour chemistry view. Part 1: Intensity and detectability. A review. Flavour and fragrance journal, 27:124-140. doi:10.1002/ffj.2090.

Ferreira, V. 2012b. Revisiting psychophysical work on the quantitative and qualitative odour properties of simple odour mixtures: A flavour chemistry view. Part 2: Qualitative aspects. A review. Flavour and fragrance journal, 27:201-215. doi:10.1002/ffj.2091.

Howard, JD & Gottfried, JA. 2014. Configural and elemental coding of natural odor mixture components in the human brain. Neuron, 84(4), 857-869. doi:10.1016/j.neuron.2014.10.012.

Hudson, R. 2000. Odor and odorant: A terminological clarification. Chemical senses, 25(6):693. doi:10.1093/chemse/25.6.693.

Jaubert, JN, Gordon, G & Dore, JC. 1987. Une organisation du champ des odeurs. Première partie : Recherche de critères objectifs. Parfums, cosmétiques, arômes, 77:53-56.

Jinks, A & Laing, DG. 2001. The analysis of odor mixtures by humans: Evidence for a configurational process. Physiology and behavior, 72(1-2):51-63. doi:10.1016/S0031-9384(00)00407-8.

Ishii, A, Roudnitzky, N, Béno, N, Bensafi, M, Hummel, T, Rouby, C & Thomas-Danguin, T. 2008. Synergy and masking in odor mixtures: An electrophysiological study of orthonasal vs. retronasal perception. Chemical senses, 33(6):553-561. doi:10.1093/chemse/bjn022.

Keller, A & Vosshall, LB. 2004. A psychophysical test of the vibration theory of olfaction. Nature neuroscience, 7(4):337-338. doi:10.1038/nn1215.

Kumar, R, Kaur, R, Auffarth, B & Bhondekar, AP. 2015. Understanding the odour spaces: A step towards solving olfactory stimulus-percept problem. Plos one, 10(10): e0141263. doi:10.1371/journal.pone.0141263.

Langlois, J, Dacremont, C, Peyron, D, Valentin, D & Dubois, D. 2011. Lexicon and types of discourse in wine expertise: The case of vin de garde. Food quality and preference, 22(6):491-498. doi:10.1016/j.foodqual.2010.10.008.

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. doi:10.1016/j.foodqual.2013.04.005.

Le Berre, E, Béno, N, Ishii, A, Chabanet, C, Etievant, P & Thomas-Danguin, T. 2008. Just noticeable differences in component concentrations modify the odor quality of a blending mixture. Chemical senses, 33:389-395. doi:10. 1093/chemse/bjn006.

Loison, A, Symoneaux, R, Deneulin, P, Thomas-Danguin, T, Fant, C, Guérin, L & Le Fur, Y. 2015. Exemplarity measurement and estimation of the level of interjudge agreement for two categories of French red wines. Food quality and preference, 40:240–251. doi:10.1016/j.foodqual.2014.10.001.

Luebke, W. 1980. The good scents company. Available at: http://www.thegoodscentscompany.com/index.html

Medjkoune, M, Harispe, S, Montmain, J, Cariou, S, Fanlo, J-L & Fiorini, N. 2016. Towards a non-oriented approach for the evaluation of odor quality. 16th International Conference, Information processing and management of uncertainty in knowledge-based systems IPMU 2016, Eindhoven, The Netherlands, 20-24 June. doi:10.1007/978-3-319-40596-4_21.

Musen, MA. 2015. The Protégé project: A look back and a look forward. AI Matters, 1(4). doi:10.1145/2557001.25757003.

Noble, AC, Arnold, RA, Buechsenstein, J, Leach, EJ, Schmidt, JO & Stern, PM. 1987. Modification of a standardized system of wine aroma terminology. American journal of enology and viticulture, 38:143-146.

Paravisini, L, Septier, C, Moretton, C, Nigay, H, Arvisenet, G, Guichard, E & Dacremont, C. 2014. Caramel odor: Contribution of volatile compounds according to their odor qualities to caramel typicality. Food research international, 57:79-88. doi:10.1016/j.foodres.2014.01.009.

Romagny, S, Coureaud, G & Thomas-Danguin, T. 2018. Key odorants or key associations? Insights into elemental and configural odour processing. Flavour and fragrance journal, 33:97-105. doi:10.1002/ffj.3429.

Rouby, C, Thomas-Danguin, T, Sicard, G, Vigouroux, M, Jiang, T, Poitevineau, J & Issanchou, S. 2005. Influence of the semantic context on odor identification performance. Psychologie française, 50:225-239. doi:10.1016/j.psfr.2004.11.003.

Snitz, K, Yablonka, A, Weiss, T, Frumin, I, Khan, RM & Sobel, N. 2013. Predicting odor perceptual similarity from odor structure. Plos computational biology, 9(9):e1003184. doi:10.1371/journal.pcbi.1003184.

Tempere, S, Hamtat, ML, Bougeant, JC, de Revel, G & Sicard, G. 2014. Learning odors: The impact of visual and olfactory mental imagery training on odor perception. Journal of sensory studies, 29(6):435-449. doi:10.1111/joss.12124.

Tempere, S, Schaaper, MH, Cuzange, E, de Lescar, R, de Revel, G & Sicard, G. 2016. The olfactory masking effect of ethylphenols: Characterization and elucidation of its origin. Food quality and preference, 50:135-144. doi:10.1016/j.foodqual.2016.02.004.

Thomas-Danguin, T, Sinding, C, Romagny, S, El Mountassir, F, Atanasova, B, Le Berre, E, Le Bon, A-M & Coureaud, G. 2014. The perception of odor objects in everyday life: A review on the processing of odor mixtures. Frontiers in psychology, 5:504. doi:10.3389/fpsyg.2014.00504.

Varela, P & Ares, G. 2012. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. Food research international, 48:893-908. doi:10.1016/j.foodres.2012.06.037.

Villière, A., Symoneaux, R., Roche, A., Eslami, A., Perrot, N., Le Fur, Y., Prost, C., Courcoux, P., Vigneau, E., Thomas-Danguin, T., Guérin, L., 2018. Dataset on the characterization of the flavor of two red wine varieties using sensory descriptive analysis, volatile organic compounds quantitative analysis by GC-MS and odorant composition by GC-MS-O. Zenodo [dataset]. doi:10.5281/zenodo.1213610.

Zarzo, M & Stanton, DT. 2006. Identification of latent variables in a semantic odor profile database using principal component analysis. Chemical senses, 31(8):713-724. doi:10.1093/chemse/bjl013.

Zarzo, M & Stanton, DT. 2009. Understanding the underlying dimensions in perfumers' odor perception space as a basis for developing meaningful odor maps. Attention, perception, & psychophysics, 71(2):225-247. doi:10.3758/APP.71.2.225.

III. Partial discussion

Within the **ARTICLE 2**, we proposed an approach based on knowledge representation techniques to formalize the odor perceptual space and the relationships among its corpuses of semantic description: odor descriptors (OD), odor qualities (OQ) and odor sensory attributes (OSA).

This new method allowed combining OD of odorants found in a food product and compiled from databases to predict an OSA profile best characterizing the whole product or odor object. The Ontology for Odor Perceptual Space (OOPS) was used as a predictive tool to estimate odor profiles of wines. In Figure 4.1, the estimated odor profiles of the two wines PN-A and CF-A using OOPS were compared with those obtained from the experimental sensory characterization of the wines (Villière et al., **Data paper included in this manuscript**). The comparison between the predicted and experimental profiles showed a good agreement between both profiles for most OSA. However, some OSA were less well predicted. The OOPS approach failed to predict the OSA Cut-grass, Leather, Toasty, Vanilla, Violet, and Woody for both wines. The OSA Violet could not be predicted for both wines and the OSA Toasty and Vanilla for the CF-A wine (set arbitrarily at the minimum value in Figure 4.1). Regarding the OSA Violet, no identified molecules in the wines carried the OD that are part of the OQ Violet, which is directly linked to the OSA Violet since Violet is a simple OSA according to expert knowledge. The same explanation applied to the OSA Toasty and Vanilla for the CF-A wine. The OSA Cut-grass, Leather and Woody are also simple OSA according to expert knowledge, which suggests that simple OSA were likely not efficiently grasped by experts. Nevertheless, further work is needed to predict the profiles for the whole set of 16 wines considered in the thesis using the OOPS approach. The expected results should help understanding why some of the OSA were less well predicted than others.

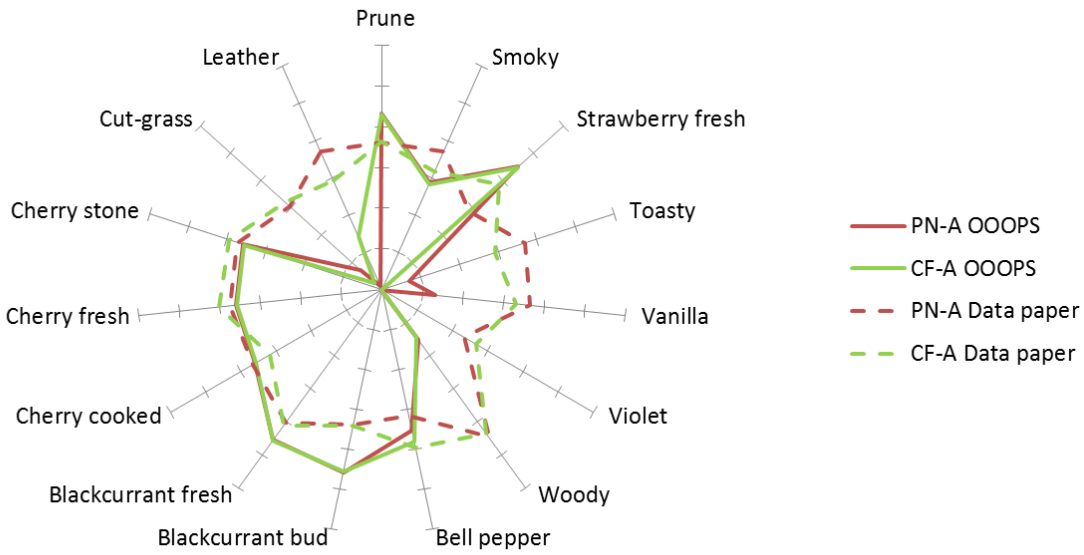


Figure 4.1: Radar plots illustrating the profile of the OSA in the PN-A and CF-A wines. Proportions of OSA are plotted in log scales (arbitrary units).

Furthermore, the ontology for odor perceptual space (OOPS) relied on expertise toward a non-targeted food product, thus the OOPS is available to study various food products in contrast with usual odor wheels which are specific to the studied product. This proof of concept should be extended to include more expertise and scientific knowledge. For instance, the sets of OD, OQ, and OSA were constructed from wine data, thus they might not cover all the possible OD, OQ, and OSA useful to characterize other food products such as cheeses or tea.

A limitation of this work was to consider the relationships among each odor space dimension as additive and linear. However, olfactory perception is inclined to be more complex as highlighted in the literature review. Indeed the odor of a mixture is most often not the sum of the odor of its components. To overcome the limitation of the OOPS, this knowledge representation relying on semantic odor information and expertise will be implemented in a broader model relying on a machine learning approach as detailed in the following chapter.

CHAPTER 5: INTEGRATIVE APPROACH BASED ON FUZZY LOGIC

I. Introduction

The chemical analysis of the odor component of food is performed by separating, identifying, and quantifying the molecules included in an extract. This well-established analytical procedure provides a list of key odorants, but does not give any information about the perceptual influence of mixed compounds, still critical to the overall food odor construction. Indeed, the odor of food is due to the processing by the olfactory system of many chemical molecules embedded in complex mixtures often recognized as single percepts due to odor blending (e.g. coffee odor). Because of the critical aspects of these perceptual integration processes, it is still very difficult to predict the odor features of a given food product on the basis of its molecular composition.

In the previous chapter, we considered experts' knowledge as a strategy to take into account the perception of odor mixture. Indeed, this knowledge was thought to fill the gap between the myriad of odor descriptors (OD) carried by odorants found in a food product and the few odor characteristics of the whole product, namely odor sensory attributes (OSA). We used an ontology as a knowledge-representation approach to formalize the experts' knowledge and used it to predict sensory outcomes, but faced the limitation that the relationships among each odor space dimension relied on additive and linear approaches.

In this chapter, we explored an innovative strategy relying on classical aroma analysis results which integrated the expert knowledge as formalized in the OOPS in a broader non-linear modelling approach relying on fuzzy logic and optimization. This strategy has been applied to predict the odor profile of the series of 16 wines and is detailed in **ARTICLE 3**.

II. ARTICLE 3

Predicting the odor profile of food from its chemical composition with an artificial intelligence modelling approach coupling fuzzy logic and expertise

Alice Roche[1], Thierry Thomas-Danguin[1], Evelyne Lutton[2], Nathalie Perrot[2]

1. CSGA, UMR1324 INRA, UMR6265 CNRS, AgroSup Dijon, Université de Bourgogne Franche - Comté, Dijon, France
2. GMPA, UMR782 INRA, AgroParisTech, Université Paris Saclay, Thiverval-Grignon, France

* Correspondence to be sent to: nathalie.perrot@inra.fr

This manuscript is in preparation.

Keywords: odor profile, machine learning, expert knowledge, fuzzy logic, prediction

Abstract

Aroma analysis follows a well-established procedure which provides a list of odorants that contribute to a given food aroma. However, such a procedure does not allow establishing the actual odor profile of the food because the perceptual influence of mixed odorants is poorly considered. To improve the aroma analysis efficiency, we explored an innovative strategy which combines aroma analysis results with an artificial intelligence modelling approach coupling expert knowledge, fuzzy logic and machine learning. The approach queries analytical and sensory databases in order to predict the odor profile of complex mixtures of odorants, namely the intensity of a series of odor sensory attributes. The developed model gathering heterogeneous data was applied to a series of real food products, i.e. wines of two grape varieties. By comparing the output of different optimization strategies and the actual sensory data we estimated that the approach can predict sensory scores in a promising way.

Introduction

Knowledge about flavor components is strategic for the whole food industry because it is a major criterion for the formulation and reformulation of food and consequently for the adaptation of food products and beverages to the increasing constraints related to nutritional, organoleptic and environmental qualities. Among the sensory dimensions involved in food flavor perception, the odor component is critical because it determines most of the time the identity and the typicality of food, which drive its overall quality and recognition by consumers (Hornung & Enns, 1986; Lorrain et al., 2006). Odors result from the perception of specific volatile molecules through the sense of smell. These specific molecules are called odorants (Hudson, 2000).

On the one hand, food odor profile can be established by sensory analysis which consists in measuring and evaluating the sensory characteristics of food through the sense of smell (Stone & Sidel, 2004). On the other hand, food odor chemical analysis is performed by separating, identifying and quantifying odorants in food products with techniques such as gas chromatography - mass spectrometry - olfactometry (GC-MS-O). This well-established analytical procedure provides a list of odorants (Lee & Noble, 2003; Ferrari et al., 2004; Pérez-Silva et al., 2006; Bratolli et al., 2013), but does not give any information about the perceptual influence of mixed odorants, still critical to the overall food odor construction.

Indeed, the odor of food is due to the processing by the olfactory system of odorants'mixtures often recognized as single percepts (e.g. coffee odor). An odor percept results from the configural processing of odor mixtures, but we are also able to discriminate odors within complex mixtures (e.g. off-odors) through the elemental processing (Thomas-Danguin et al., 2014). Because of the critical aspects of these perceptual integration processes, it is still very difficult to predict the odor profile of a given food product on the basis of its chemical composition. Within the flavor analysis path, additional time-consuming omission tests are often required to ascertain the role of each odorant and/or their mixtures in the overall flavor (Grosh, 2001 / wine: Guth, 1997; Ferreira et al., 2002; Escudero et al., 2004; Ferreira et al., 2016 / sesame: Zheng et al., 2016).

Predictive approaches based on the molecular structure of odorants were developed to predict the odor characteristics of new sets of odorants (intensity: Cain, 1969; Edwards & Jurs, 1989 / detection threshold: Abraham et al., 2012 / perceptual qualities: Kumar et al., 2015 / pleasantness: Khan et al., 2007). Other strategies aimed to develop machine learning algorithms to predict several odor characteristics of odorants based on their structural parameters (Keller et al., 2017, Shang et al., 2017). Although these predictive approaches were successful, they all applied to single odorants. However, the odors we perceive in every-day life are complex multi-molecular mixtures. Snitz et al. (2013) worked on the odor characteristics of such mixtures by predicting the similarity of multi-molecular mixtures on the basis of their structural parameters. However, to our knowledge no study has attempted to predict how a multi-molecular mixture will smell, namely to generate a quantitative sensory description of the mixtures' odor using semantic descriptors.

Because classical/mechanistic approaches (e.g. summation) are not sufficient to study both elemental and configural perceptions of odorants' mixtures, we chose to integrate expertise from flavorists in a highly innovative predictive model. Indeed, the way flavorists combine odor qualities (e.g. Fruity, Green, Smoky) or molecules to create a target odor represents their expertise. Such expertise relied on a very personal process (Veinand, 2015) and at the same time is based on generic principles learnt in specialized schools (e.g. based on the field of odors described by Jaubert et al., 1995) or companies (e.g. Sense It^{TM}, a global flavor language at Givaudan). Collecting this expertise was thought to better understand how flavor are created and thus how an odor profile is constructed.

Expertise can be formalized in the form of an in-silico model, using fuzzy logic (Sicard et al., 2011), which is an extension of the set theory by the replacement of the characteristic function of a set by a membership function whose values range from 0 to 1 (Zadeh, 1965). Transitions between sets allow the representation of gradual concepts as well as the representation of rules, particularly adapted to cope with applications where expertise is present (Perrot et al., 2006). Analogies exist between fuzzy entities and sensory entities (i.e. sensory scales as fuzzy sets, sensory attributes as fuzzy variables, and sensory answers as membership grades) (Tan et al., 1999; Davidson et al., 2001), which explains why fuzzy logic has been used successfully in several papers dealing with food systems (Perrot et al., 2011). We can cite its application to estimate the sensory properties of food products (sausage: Ioannou et al., 2002 / chhana podo: Mukhopadhyay et al., 2013 / tea liquor: Debjani et al., 2013 / jam: Shinde & Pardeshi, 2014) and predict consumer food acceptance (biscuits: Davidson et al., 2001 / bread: Folorunso et al., 2009).

The challenge of the present research was to increase the efficiency of the odor analysis procedure by taking into account knowledge on odor mixture perception along the analytical path. In this proof-of-concept study, we tested the possibility that a machine learning based prediction model might be used to predict the odor profile of 16 red wines on the basis of their chemical composition. To construct this predictive model we integrated flavorists' expertise, gathered in an ontology format (Roche et al., **Article 2 of this manuscript**) and artificial intelligence methodologies, i.e. fuzzy logic modelling and genetic algorithm's based optimization; to formalize the expertise in a set of optimized rules of sensory relevant odor combinations.

Materials & Methods

Chemical and sensory data on wines

Chemical and sensory characterizations of the wines were published in a data paper (Villière et al., **Data paper included in this manuscript**). Gas chromatography - olfactometry (GC-O) and gas chromatography - mass spectrometry (GC-MS) results were obtained for a set of 16 French red wines, 8 from Pinot Noir (PN) and 8 from Cabernet Franc (CF) selected in order to cover the olfactory diversity of the commercial offer for each grape variety (Loison et al.,

2015). GC-MS-O data provided the identification of a total of 46 odorant zones. GC-O technique assigns a relative importance to each odor-active compound by calculating their nasal impact frequency (NIF), which is the ratio of the number of panelists who smelled an odor on the number total of panelists (Pollien et al., 1997). The wines were submitted to an expert sensory panel to construct odor profiles.

Odor description of the wines' odorants

We compiled three databases to collect the odor descriptors of the odorants identified by in the 16 wines: Arctander's handbook (Arctander, 1969), Flavor-Base (Leffingwell & Associates, http://www.leffingwell.com) and The good scents company (Luebke, 1980). This step followed the methodology described in (Roche et al., **Article 2 of this manuscript**).

Expert knowledge

Four senior flavorists participated in the data collection. The elicitation process was based on a private guided interview which lasts 1 hour. The experts received monadically 15 odor sensory attributes (OSA) of interest to characterize the wines but were not informed that the descriptors were related to wines. For each attribute, they were asked to indicate the odor notes, further called odor qualities (OQ), needed to construct a given OSA. The collected knowledge was gathered in the form of an ontology (Roche et al., **Article 2 of this manuscript**). Proportions of OQ were calculated from the collected data, the sum of OQ for a given expert and OSA being equal to 1. One OSA can be composed of one OQ (simple OSA) or several OQ (complex OSA). Expert knowledge was formalized in a set of optimized rules of sensory relevant odor combinations using fuzzy logic.

Modelling and fuzzification

The expert model was developed on MATLAB (R2014b), using linear regression for the simple OSA and fuzzy reasoning (Zadeh, 1975; Dubois & Prade, 1980) for the complex OSA, in order to estimate the intensity of each OSA for a food product on the basis of its proportion in OQ.

The fuzzy reasoning parameters, which are defining the membership functions, were established based on the expert knowledge we collected. Considering an OSA described by the experts as a complex OSA composed of five odor qualities, $OQi\ _{i=1}^{5}$, with different proportions depending on the expert. A fuzzy set E in universe of discourse U can be defined for each odor quality by Equation 1 (Zadeh, 1965):

Equation 1:

$$E = \{u, \mu_E(u) \backslash u \in U\}$$

$$\mu_E : U \rightarrow [0,1]$$

$\mu_E(u)$ is the membership function of the set E. It represents the set of membership grades $\mu_E(u)$ of a numerical variable u (OQi) mapped to a fuzzy set E (OSA). The value of the membership grade is a real number within the interval [0,1], representing the membership degree of the proportion of each OQi to a given OSA intensity. The proportions were defined by the expert flavorists. Membership functions can be expressed through various representations and we chose to test our model with the trapezoidal representations using four parameters a1 to a4 (Equation 2), represented graphically in Figure 1.

Equation 2:

$$\mu(x) = \begin{cases} 0\ (x \leq a_1) \\ \frac{x-a_1}{a_2-a_1}\ (a_1 < x \leq a_2) \\ 1\ (a_2 < x \leq a_3) \\ \frac{a_4-x}{a_4-a_3}\ (a_3 < x \leq a_4) \\ 0\ (a_4 < x) \end{cases}$$

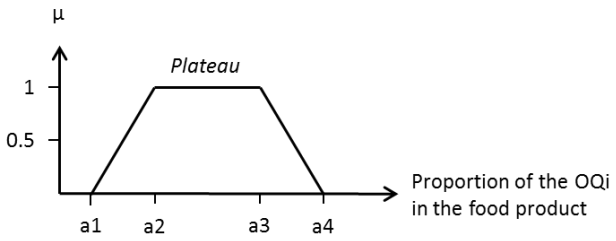


Figure 1: Fuzzy trapezoidal function. a2 and a3 represent the bounds of the proportion of the OQi when the membership degree to a given OSA equals to 1. a1 varies between 0 and a2, a4 varies between a3 and 1.

The prediction of OSA is achieved using a generalization of the inference compositional rule proposed by Zadeh (1975) applied to the fuzzy symbolic descriptions (Mauris et al., 1994) of OQ. Mathematical product is used for the T-norm (Equation 3).

Equation 3: $$Intensity(OSA) = 10 \times \prod_1^i \mu(OQ_i)$$

Optimization

The parameters (a1 to a4) of the fuzzy membership functions were optimized using a genetic algorithm (condition B) and an evolutionary covariance matrix adaptation evolutionary algorithm (condition C).

The genetic algorithm (repetitions = 10) was set to minimize the difference between the sensory evaluated and sensory predicted values (cost function = fuzzy function, maximum number of iterations = 500, population size = 200, offspring size = 200, crossover percentage = 0.7, mutation percentage = 0.3, tournament size = 3).

The evolutionary covariance matrix adaptation evolutionary algorithm was adapted from cmaes.m (Version 3.61.beta) and ran for 10 repetitions (cost function = fuzzy function, population size = 2000).

Validation

The expert model was validated following a leave-one-out cross-validation (LOOCV) procedure, meaning that each sample (food product) is left out once and used for validation (Arlot & Celisse, 2010).

The outputs of the fuzzy models were the intensity predicted for the 15 OSA. We compared these predicted intensities with the intensity evaluated from sensory evaluation. First, we calculated the percentage prediction error following Equation 4. The closer the sensory intensity evaluated and measured the lower percentage prediction error. In contrary, if the sensory intensity evaluated and measured are highly different, the percentage will be high (100% and more) (Guang et al., 1995).

Equation 4: $$Percentage\ predicted\ error = \frac{|evaluated - predicted|}{evaluated} \times 100$$

In addition, we calculated the similarity (sim) between the intensities evaluated and predicted for each OSA following the Ruzicka similarity (Equation 5). The values vary from 0 (no similarity) to 1 (datasets identical).

Equation 5: $$sim(X, Y) = \frac{\sum min\{x_i, y_i\}}{\sum max\{x_i, y_i\}}$$

Finally, we performed a normalized principal component analysis (PCA) and a hierarchical clustering on principle components (HCPC) on the sensory profile from sensory evaluation and from the predictive approach.

Results

The predictive approach

The developed predictive approach gathered different types of data and models (Figure 2). The data were the results of chemical analysis, expertise data collection and sensory analysis. Concerning the models, we used the Ontology for Odor Perceptual Space (OOPS) (Roche et al., **Article 2 of this manuscript**) and we constructed an artificial intelligence model to predict the odor profiles of a set of 16 red wines. First, we considered the chemical composition of the wines in terms of identified odorants and their NIF (Villière et al., **Data paper included in this manuscript**). Then, we aggregated the information of three databases to obtain the odor descriptors of the odorants identified in the wines. For a given odorant, a set of odor descriptors was obtained from the database aggregation in which each term was associated to an intensity (number of citation of a given odor descriptors in the three databases). For example, the set of odor descriptors of Isoamyl acetate (CAS 123-92-2) is the following: [(apple, 1); (banana, 3); (fruity, 2); (fruity-fresh, 1); (nauseating, 1); (pear, 2); (solvent, 1); (sweet, 3)]. Then we weighted each set of odor descriptors by the relative importance of the corresponding odorant in a given food product (i.e. its NIF score). Thus, the proportions of odor descriptors of Isoamyl acetate in a given food product with a NIF equals to NIF1 will be [(apple, 1xNIF1); (banana, 3xNIF1); (fruity, 2xNIF1); (fruity-fresh, 1xNIF1); (nauseating, 1xNIF1); (pear, 2xNIF1); (solvent, 1xNIF1); (sweet, 3xNIF1)]. This weighted set is translated into a set of odor qualities (OQ) thanks to the OOPS (Roche et al., **Article 2 of this manuscript**), which results in: [(Fresh, 1xNIF1); (Fruity, 9xNIF1)]. For a given food product, the OQ of each odorant identified in the product were summed. For instance, if one consider a mixture of Isoamyl acetate and Benzyl alcohol (CAS 100-51-6), with NIF scores respectively NIF1 and NIF2, it is possible to sum the OQ set of both odorants [(Fresh, 1xNIF1); (Fruity, 9xNIF1)] and [(Almond, 1xNIF2); (Floral, 2xNIF2); (Fruity, 1xNIF2); (Smoky, 1xNIF2)], to end up with an OQ set being [(Almond, 1xNIF2); (Floral, 2xNIF2); (Fresh, 1xNIF1); (Fruity, 9xNIF1+1xNIF2); (Smoky, 1xNIF2)]. Then, we processed the data through an artificial-intelligence model based on flavorists' expertise. Expertise was formalized and coded in fuzzy rules in order to estimate the odor sensory attributes (OSA) intensity of the wines from their OQ sets. The output of the predictive model was the odor sensory profile of a wine. The predicted profile was compared to the odor sensory profile of the same wines obtained by sensory evaluation.

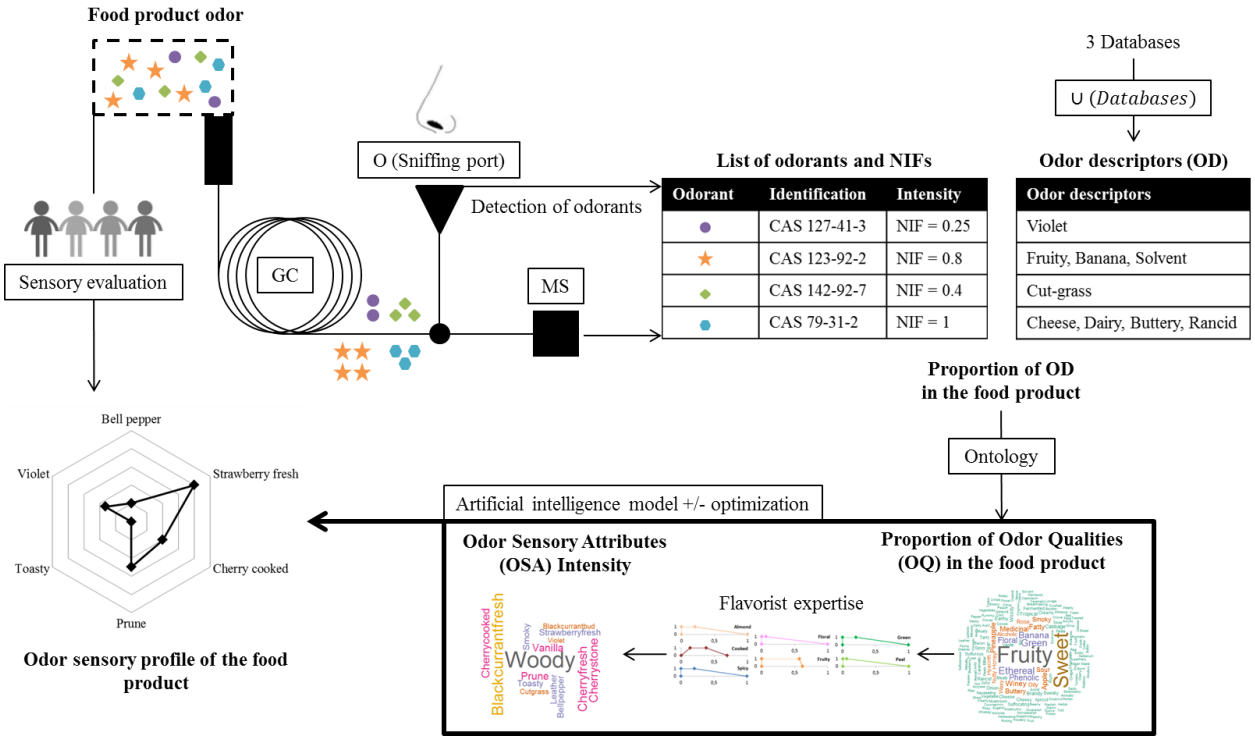


Figure 2: Schematization of the predictive approach. The boxes represent the methods and arrows are pointing to the results obtained through the given method. On the one hand, the food product odor is analyzed by GC-MS-O (gas chromatography - mass spectrometry - olfactometry) which allows the identification of odorants and their intensity with their nasal impact frequency (NIF) scores. On the other hand, odor descriptors are compiled from the union of three databases (Arctander, Flavor-Base and The good scents company). From these two results, we obtain the proportion of odor descriptors in the food product which is then processed through the Ontology for Odor Perceptual Space (OOPS; Roche et al., **Article 2 of this manuscript**) to be translated in proportion of odor qualities (OQ). At this step, the artificial intelligence model computes the proportion of OQ in the food product owing to flavorists' expertise to estimate the intensity of odor sensory attributes (OSA). The output of the model is the odor sensory profile of the food product which can be compared with the sensory profile obtained by sensory evaluation.

Artificial intelligence model

Expertise

Expert flavorists were asked to decompose 15 OSA (Table 1) in one or several OQ, i.e. simple or complex OSA. Usually flavorists have in mind a way to combine OQ to elicit the perception of a given OSA in complex mixtures and matrices. Here, their reasoning is formalized in a backwards way using fuzzy functions and fuzzy rules establishing a projection from the OQ space to the OSA space. Among the 15 OSA, 7 were identified as simple OSA (Cut-grass, Leather, Smoky, Toasty, Vanilla, Violet and Woody) and 8 as complex OSA (Bell pepper, Blackcurrant bud, Blackcurrant fresh, Cherry cooked, Cherry fresh, Cherry stone, Prune and Strawberry fresh). The detailed results of the interview for the OSA Prune (complex OSA) are presented below in Table 2. The compiled results of the interview (i.e. the proportions bounds: lower and higher proportions) for the 15 OSA are presented in Table 3.

Table 1: List of the 15 odor sensory attributes (OSA).

Bell pepper	Prune
Blackcurrant bud	Smoky
Blackcurrant fresh	Strawberry fresh
Cherry cooked	Toasty
Cherry fresh	Vanilla
Cherry stone	Violet
Cut-grass	Woody
Leather	

Table 2: Composition of the OSA Prune in OQ from expertise. The proportions of each OQ composing the OSA Prune are indicated for each expert into symbolic scale (+++, ++, +, +/-, near 0) translated into numerical values. *Nu* means that the OQ is not used by the given expert. Values in **bold** represent the lower and higher proportion of each odor quality composing the OSA.

Prune	Almond	Cooked	Fruity	Honey	Lactonic
Expert 1	+	*Nu*	*Nu*	*Nu*	++
	0.33	*Nu*	*Nu*	*Nu*	0.67
Expert 2	*Nu*	++	+	near 0	+
	Nu	0.47	0.24	0.06	0.24
Expert 3	*Nu*	*Nu*	*Nu*	*Nu*	*Nu*
Expert 4	*Nu*	+++	+	*Nu*	*Nu*
	Nu	0.75	0.25	*Nu*	*Nu*
Lower proportion	**0**	**0**	**0**	**0**	**0**
Higher proportion	**0.33**	**0.75**	**0.25**	**0.06**	**0.67**

Table 3: Composition of the 15 OSA in OQ from expertise: 7 simple OSA and 8 complex OSA. The proportions of each OQ composing the complex OSA are represented through their lower and higher proportions from the expertise data.

OSA	OQ1	OQ2	OQ3	OQ4	OQ5	OQ6	OQ7
Cut-grass	Cut-grass						
Leather	Leather						
Smoky	Smoky						
Toasty	Toasty						
Vanilla	Vanilla						
Violet	Violet						
Woody	Woody						
Bell pepper	Floral	Fruity	Green	Sulfurous	Toasty	Vegetable	
Lower proportion	0	0	0	0.17	0	0	
Higher proportion	0.17	0.33	0.80	0.25	0.5	0.33	
Blackcurrant bud	Floral	Fresh	Fruity	Green	Sulfurous	Vanilla	Wine-like
Lower proportion	0	0	0.10	0	0.11	0	0
Higher proportion	0.22	0.33	0.67	0.20	0.40	0.20	0.38
Blackcurrant fresh	Floral	Fresh	Fruity	Green	Sulfurous	Wine-like	
Lower proportion	0	0	0.25	0	0.07	0	
Higher proportion	0.13	0.53	0.44	0.22	0.25	0.25	
Cherry cooked	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
Lower proportion	0.13	0.14	0	0	0	0	0
Higher proportion	0.29	0.38	0.07	0.57	0.21	0.06	0.21
Cherry fresh	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
Lower proportion	0.12	0.22	0	0	0.22	0	0
Higher proportion	0.26	0.24	0.07	0.35	0.24	0.06	0.22
Cherry stone	Almond	Cooked	Floral	Fruity	Green	Peel	Spicy
Lower proportion	0.25	0	0	0	0	0	0
Higher proportion	1	0.25	0.06	0.38	0.18	0.06	0.18
Prune	Almond	Cooked	Fruity	Honey	Lactonic		
Lower proportion	0	0	0	0	0		
Higher proportion	0.33	0.75	0.25	0.06	0.67		
Strawberry fresh	Cooked	Floral	Fruity	Green			
Lower proportion	0.17	0	0.33	0			
Higher proportion	0.40	0.33	0.5	0.27			

Fuzzification and optimization

In order to explain the general approach, we will keep the example of the OSA Prune, described by the experts as a complex OSA composed of five OQ (Almond, Cooked, Fruity, Honey, Lactonic) with different proportions depending on the expert. Five fuzzy membership functions, one per OQ, were created from expertise (Figure 3). The functions were used to link the proportion of the OQ to the intensity of the OSA.

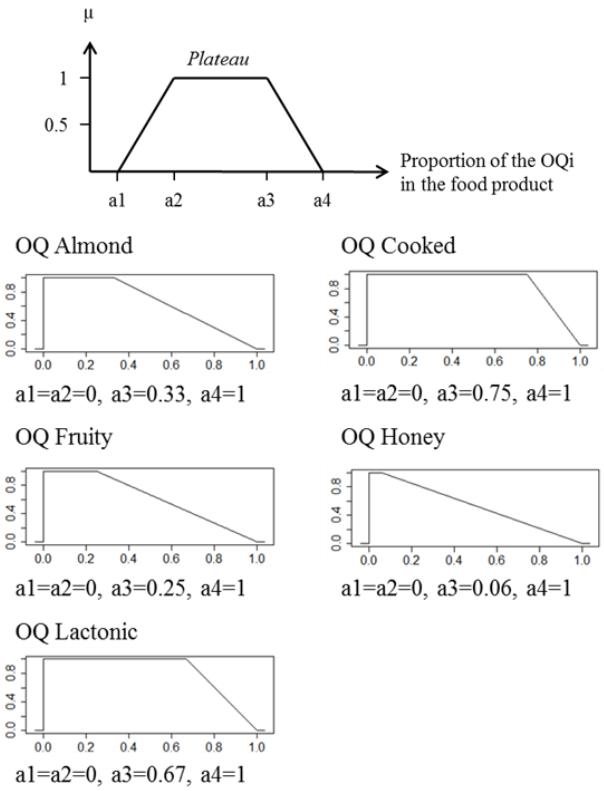


Figure 3: Fuzzy membership functions created for the OSA Prune for the non-optimized condition A. a2 and a3 represent the bounds of the proportion of the OQ when the membership degree to a given OSA equals to 1 (lower and higher proportion). a1 varies between 0 and a2 while a4 varies between a3 and 1, in this example we set a1 = 0 and a4 = 1 for the five functions.

Fuzzy rules were then computed following the Equation 6, meaning that if the membership degree of the five OQ to the intensity of the OSA Prune is equal to 1 (plateau), the intensity of the OSA Prune will be 10/10.

Equation 6: $$Intensity(OSA\ Prune) = 10 \times \prod_1^5 \mu(OQ_i)$$

with $OQ_1 = Almond, OQ_2 = Cooked, OQ_3 = Fruity, OQ_4 = Honey\ and\ OQ_5 = Lactony$

The parameters (a1 to a4) of the fuzzy membership functions were non-optimized or optimized for each OQ composing an OSA according to three conditions as described in the Table 4. The non-optimized condition (condition A) corresponded to the creation of the fuzzy membership functions in Figure 3. Then, the optimization of the parameters was done in order to minimize the gap between the sensory values (OSA intensity evaluated by sensory evaluation) and the predicted values (OSA intensity predicted by the model). We tested a first optimization based on a genetic algorithm (GA) (Mitchell, 1996) and a second based on a covariance matrix adaptation evolutionary algorithm (Hansen, 2006). The first one aimed to optimize 2 out of the 4 parameters and the second one optimized the 4 parameters of the membership functions. In both cases, 10 repetitions were carried out and we further considered the mean of the intensity predicted for these 10 repetitions as the intensity predicted for a food product.

Table 4: Non-optimized and optimized parameters of the fuzzy memberships functions according to three conditions. Proportions of OQ are from the expert knowledge collection.

	Optimization	a1	a2	a3	a4
Condition A	Non-optimized 0-1	0	Lower proportion of the OQ	Higher proportion of the OQ	1
Condition B	Optimized genetic algorithm	optimized	Lower proportion of the OQ	Higher proportion of the OQ	optimized
Condition C	Optimized cmaes	optimized	optimized	optimized	optimized

Predicting the intensity of OSA: illustrative example for Toasty and Prune

We chose to present the predictive approach in details to predict the intensity of one simple OSA (Toasty) and one complex OSA (Prune) for the wine PN1, a Pinot Noir wine out of the 16 wines of the sample set. The list of the 33 odorants identified in PN1 with their corresponding NIF scores and odor descriptors sets are shown in Table 5. The odor descriptors were translated into OQ thanks to the OOPS (Roche et al., **Article 2 of this manuscript**) and weighted by the NIF scores of the corresponding odorant.

The odor qualities set of the 33 odorants were summed to obtain the odor qualities set of PN1: (Almond, 0.625); (Cooked, 2.375); (Cut-grass, 1.5); (Floral, 16.875); (Fresh, 0.125); (Fruity, 58.625); (Green, 8.625); (Honey, 2.5); (Lactonic, 0.5); (Leather, 2.25); (Peel, 3.625); (Smoky, 18.75); (Spicy, 6.375); (Sulfurous, 0.875); (Toasty, 1); (Vanilla, 2.625); (Vegetable, 3.5); (Violet, 0); (Wine-like, 4.25); (Woody, 3.25). This set was further represented as proportions of OQ, following $prop(Almond) = \frac{intensity(Almond)}{\sum intensity(20\ OQ)} = \frac{0.625}{138.3} = 0.005$, to result in the odor quality set (Almond, 0.005); (Cooked, 0.017); (Cut-grass, 0.011); (Floral, 0.122); (Fresh, 0.001); (Fruity, 0.424); (Green, 0.062); (Honey, 0.018); (Lactonic, 0.004); (Leather, 0.016); (Peel, 0.026); (Smoky, 0.136); (Spicy, 0.046); (Sulfurous, 0.006); (Toasty, 0.007); (Vanilla, 0.019); (Vegetable, 0.025); (Violet, 0); (Wine-like, 0.031); (Woody, 0.024).

Then we processed the odor quality set through the fuzzy model to determine the intensity of the OSA Toasty and Prune. We predicted the OSA intensity owing to a leave one out cross validation, so that the model was constructed on a set of 15 wines to predict the omitted one. The intensity predicted from the model was then compared to the intensity evaluated from sensory evaluation.

Table 5: List of odorants in the wine PN1. Odorants are identified by their CAS number, NIF scores, set of odor descriptors and set of odor qualities. NIF scores correspond to the ratio of the number of panelists who perceived the odorant on the number total of panelists (n = 8). The odor descriptors sets are obtained after compiling three databases (Arctander, Flavor-Base, The good scents company). The odor qualities set are obtained following the Ontology for Odor Perceptual Space (OOPS) approach and are weighted with the NIF of the corresponding odorant.

CAS Number	NIF	Odor descriptors	Odor qualities
100-51-6	0.625	(almond, 1); (balsamic, 1); (floral, 1); (fruity, 1); (phenolic, 1); (rose, 1); (sweet, 2)	(Almond, 0.625); (Floral, 1.25); (Fruity, 0.625);(Smoky, 0.625)
105-54-4	0.875	(banana, 2); (buttery, 1); (cognac, 1); (ethereal, 1); (ethereal-fruity, 1); (fruity, 2); (juicy, 2); (pineapple, 3); (ripe fruit, 1)	(Fruity, 7.875)
106-32-1	0.125	(apricot, 2); (banana, 2); (brandy, 1); (fermented-winey, 1); (fruity, 2); (fruity-winey, 1); (pear, 1); (pineapple, 1); (sweet, 3); (waxy, 1); (winey, 2)	(Fruity, 1.125); (Wine-like, 0.5)
106-33-2	0.5	(fatty, 1); (floral, 2); (flower-petal, 1); (fruity, 2); (leafy, 1); (oily, 1); (oily-fatty, 1); (soapy, 1); (sweet, 1); (waxy, 2)	(Floral, 1.5); (Fruity, 1)
106-44-5	0.5	(animal, 1); (animalic, 1); (dry, 1); (dry-tarry, 1); (leather, 1); (leathery, 1); (medicinal, 3); (medicinal-leathery, 1); (mimosa, 1); (narcissus, 1); (phenolic, 3); (smoky, 1); (tarry, 1); (tarry-smoky, 1); (woody, 1)	(Floral, 1); (Leather, 1.5); (Smoky, 5); (Woody, 0.5)
107-92-6	0.625	(acetic, 1); (buttery, 2); (cheese, 1); (cheesy, 1); (fruity, 1); (rancid, 1); (rancid butter, 1); (sour, 2)	(Fruity, 0.625)
108-64-5	0.875	(apple, 3); (banana, 1); (blueberry, 1); (buttery, 1); (ethereal, 1); (fruity, 2); (pineapple, 1); (sweet, 2); (tutti fruit, 1); (wine-like-fruity, 1); (winey, 1)	(Fruity, 8.75); (Wine-like, 1.75)
108-95-2	0.125	(medicinal, 1); (phenolic, 2); (plastic, 1); (rubbery, 1)	(Smoky, 0.25)
122-78-1	0.75	(clover, 1); (cocoa, 1); (floral, 3); (green, 3); (honey, 2); (hyacinth, 3); (rose, 1); (sweet, 2)	(Floral, 5.25); (Green, 2.25); (Honey, 1.5)
123-07-9	0.75	(castoreum, 1); (guaiacol, 1); (phenolic, 1); (smokey, 1); (smoky, 1); (sweet, 1); (tarry-medicinal, 1)	(Leather, 0.75); (Smoky, 3.75); (Spicy, 0.75)
123-51-3	1	(alcoholic, 3); (banana, 1); (fermented, 1); (fruity, 1); (fruity-winey, 1); (fusel, 1); (whiskey, 1)	(Fruity, 3); (Wine-like, 1)
123-66-0	1	(apple peels, 1); (banana, 3); (floral, 1); (fruity, 2); (fruity-winey, 1); (green, 1); (pear, 1); (pineapple, 3); (strawberry, 1); (sweet, 1); (tropical, 1); (waxy, 1)	(Floral, 1); (Fruity, 13); (Green, 1); (Peel, 1); (Wine-like, 1)
123-92-2	0.125	(apple, 1); (banana, 3); (fruity, 2); (fruity-fresh, 1); (nauseating, 1); (pear, 2); (solvent, 1); (sweet, 3)	(Fresh, 0.125); (Fruity, 1.125)

CAS Number	NIF	Odor descriptors	Odor qualities
2785-89-9	0.625	(bacon, 2); (clove, 2); (eugenol, 1); (guaiacol, 1); (phenolic, 2); (smoky, 2); (spicy, 2); (spicy-medicinal, 1); (sweet, 1); (vanilla, 1)	(Smoky, 3.125); (Spicy, 4.375); (Vanilla, 0.625)
334-48-5	0.375	(cheese, 1); (citrus, 1); (dairy, 1); (fatty, 2); (rancid, 2); (sour, 1); (sour-fatty, 1); (waxy, 1)	(Fruity, 0.375)
3268-49-3	0.625	(bouillon, 1); (creamy, 1); (earthy, 1); (meaty, 1); (musty, 1); (onion, 2); (onion-meat, 1); (potato, 1); (tomato, 1); (vegetable, 1)	(Vegetable, 1.875)
39212-23-2	0.5	(burnt, 1); (celery, 1); (coconut, 2); (coumarinic, 2); (lactonic, 1); (lovage, 1); (maple, 1); (nutty, 1); (toasted, 1); (tonka, 1); (woody, 1)	(Lactonic, 0.5); (Spicy, 0.5); (Toasty, 1); (Vanilla, 0.5); (Vegetable, 0.5); (Woody, 0.5)
431-03-8	0.875	(buttery, 3); (caramel, 1); (chlorine-quinone, 1); (creamy, 1); (oily, 1); (sweet, 1)	(Cooked, 0.875)
4312-99-6	0.5	(earthy, 1); (herbal, 1); (metallic, 1); (mushroom, 2); (musty, 1)	(Cut-grass, 0.5); (Green, 0.5)
503-74-2	1	(acid-acrid, 1); (cheese, 1); (cheesy, 2); (fruity, 1); (herbaceous, 1); (sour, 2); (sweaty, 2); (tropical, 1)	(Cut-grass, 1); (Fruity, 2); (Green, 1)
505-10-2	0.25	(meaty, 2); (mushroom, 1); (onion, 1); (soup, 2); (sulfuraceous, 1); (sulfurous, 1); (sweet, 1); (sweet soup-meat, 1); (vegetable, 1)	(Sulfurous, 0.5); (Vegetable, 0.75)
590-86-3	0.25	(acrid, 1); (aldehydic, 1); (cheese, 1); (chocolate, 1); (cocoa, 1); (ethereal, 1); (fatty, 1); (fruity, 2); (green fruity, 1); (peach, 2); (sweaty, 1)	(Fruity, 1.25); (Green, 0.25)
60-12-8	1	(earthy, 1); (floral, 2); (greener gassy, 1); (hyacinth, 1); (rose, 2); (rose-honey, 1)	(Floral, 6); (Green, 1); (Honey, 1)
64-17-5	0.625	(alcoholic, 3); (ethereal, 2); (medicinal, 1); (sweet, 1); (sweet-ethereal, 1)	(Fruity, 0.625)
64-19-7	1	(acidic, 1); (sour, 3); (vinegar, 2)	
620-17-7	0.75	(medicinal, 1); (musty, 1); (sweet, 1); (woody-phenolic, 1)	(Smoky, 0.75); (Woody, 0.75)
74-93-1	0.375	(cabbage, 1); (garlic, 1); (rotten cabbage, 1); (rotting cabbage, 1); (sulfurous, 1)	(Sulfurous, 0.375); (Vegetable, 0.375)
7452-79-1	0.875	(apple, 2); (apple peels, 1); (fruity, 3); (green, 2); (green-fruity, 1); (peels of unripe plums, 1); (pineapple skin, 1); (strawberry, 1); (sweet, 1)	(Fruity, 0.875); (Green, 2.625); (Peel, 2.625)
80-62-6	0.625	(acrylate, 1); (acrylic, 1); (apple, 1); (estery, 1); (fruity, 2); (grape, 1)	(Fruity, 3.125)
90-05-1	0.75	(medicinal, 2); (phenolic, 2); (smoky, 3); (spicy, 1); (sweet, 1); (vanilla, 2); (woody, 1)	(Smoky, 3.75); (Spicy, 0.75); (Vanilla, 1.5); (Woody, 0.75)
91-10-1	0.375	(bacon, 1); (balsamic, 1); (medicinal, 1); (phenolic, 2); (powdery, 1); (smoky, 2); (woody, 2)	(Smoky, 1.5); (Woody, 0.75)
96-48-0	0.75	(buttery, 1); (caramel, 2); (creamy, 1); (fatty, 1); (nutty, 1); (oily, 1); (sweet, 1)	(Cooked, 1.5)
97-62-1	0.875	(alcoholic, 1); (apple, 1); (ethereal, 2); (floral, 1); (fruity, 3); (fusel, 1); (rum, 1); (rummy, 1); (sweet, 3); (sweet-ethereal, 1)	(Floral, 0.875); (Fruity, 4.375)

Predicted intensity of a simple OSA: Toasty

For simple OSA, such as Toasty, a linear regression was performed between the proportion of the OQ Toasty and the intensity of the OSA Toasty for the set of 15 wines (16-1 for the cross validation). We obtained the linear equation $y = 29.67x + 0.37$ ($r^2 = 0.2694$, $p = 0.03$) which allowed the prediction of the intensity of the OSA Toasty in a food product from the proportion of the OQ Toasty in the product. Knowing that the proportion of the OQ Toasty is 0.008 in the wine to predict (PN1), the intensity predicted for the OSA Toasty for PN1 was 0.60 (to be compared to the actual intensity evaluated by sensory analysis = 0.61).

Predicted intensity of a complex OSA: Prune

Focusing on the specific OSA Prune, the OQ set was reduced to the proportions of the 5 OQ constituting this OSA. In this way, from the odor quality set of PN1 (Almond, 0.005); (Cooked, 0.017); (Cut-grass, 0.011); (Floral, 0.122); (Fresh, 0.001); (Fruity, 0.424); (Green, 0.062); (Honey, 0.018); (Lactonic, 0.004); (Leather, 0.016); (Peel, 0.026); (Smoky, 0.136); (Spicy, 0.046); (Sulfurous, 0.006); (Toasty, 0.007); (Vanilla, 0.019); (Vegetable, 0.025); (Violet, 0); (Wine-like, 0.031); (Woody, 0.024), we kept the values corresponding to the OQ of the OSA Prune: (Almond, 0.005); (Cooked, 0.017); (Fruity, 0.424); (Honey, 0.018); (Lactonic, 0.004). The values were then converted in order to have the sum of the 5 OQ being equal to 1, we ended up with (Almond, 0.01); (Cooked, 0.037); (Fruity, 0.907); (Honey, 0.039); (Lactonic, 0.008).

In this case, the relationship between the chemical composition of the wine and the sensory data was modeled through fuzzy rules according to three conditions of optimization. Fuzzy membership functions linking the proportion of the five OQ Almond, Cooked, Floral, Fruity, Green, Peel, Spicy and the intensity of the OSA Prune in 15 wines (16-1 for the LOOCV) according to the three optimization conditions are shown in Figure 4 where the proportions of the five OQ Almond, Cooked, Fruity, Honey and Lactonic in the wine PN1 are represented as red triangles. Parameters of the fuzzy functions (a1 to a4) are represented in Table S2.

Then we calculated the intensity predicted for the OSA Prune in the wine PN1 for the three optimization conditions, following Equation 3. The intensity evaluated by sensory analysis for the OSA Prune in the wine PN1 was 0.95. With regard to the predicted values, two of the three optimization conditions (condition A and condition B) did perform better by predicting

an intensity of this OSA being 1.20 compared to condition C which predicted an intensity of 1.62.

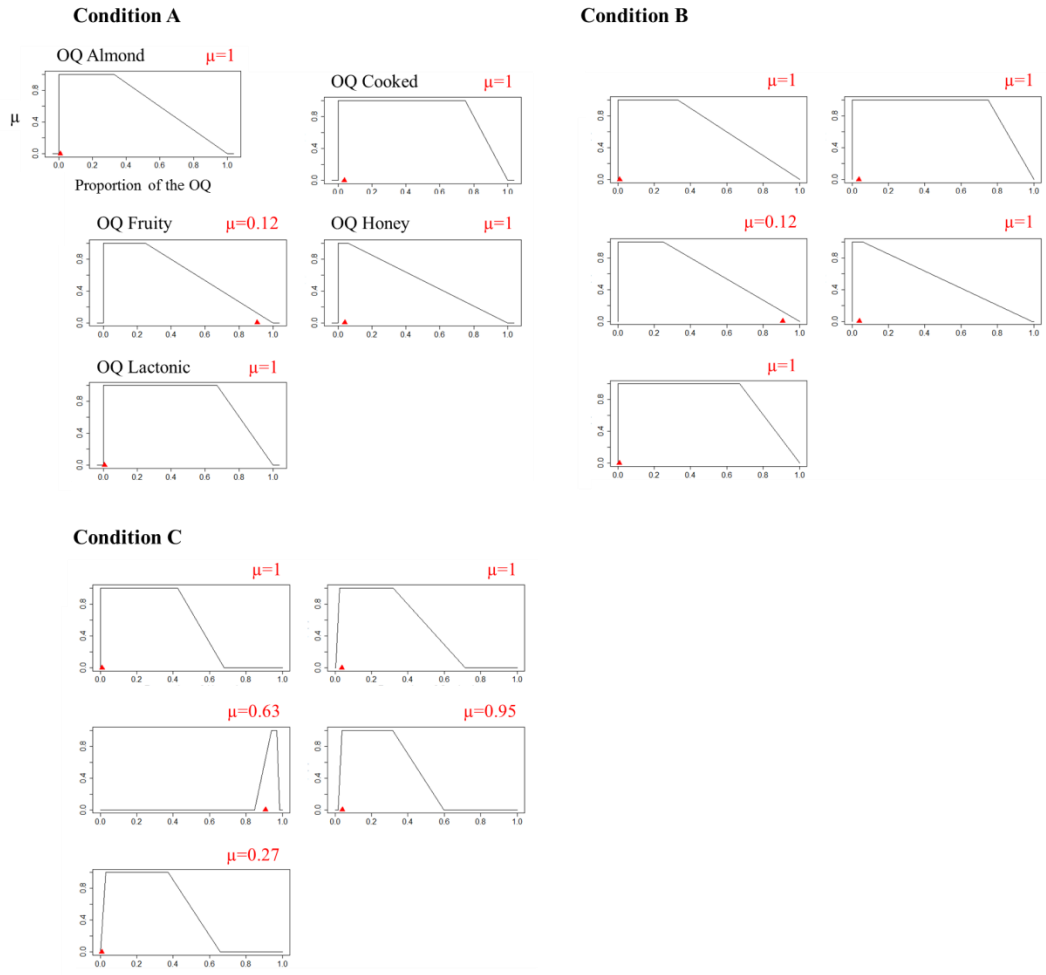


Figure 4: Fuzzy membership functions linking the proportion of the 5 OQ Almond, Cooked, Fruity, Honey and Lactonic to the intensity of the OSA Prune according to the three conditions of optimization. Red triangles are representing the proportions of the OQ in the wine to predict (PN1).

Predicting the odor sensory profiles of wine and validation

The predictive approach was applied to the 16 wines of the sample set. The goal was to predict the intensity of 15 OSA for the 16 wines. The quality of the prediction was estimated with linear regression between the measured and predicted intensity of the OSA. Figure S1 illustrated the regressions of the 7 simple OSA and Figure S2 the 8 complex OSA for the

three conditions of optimization. In S3 and S4 Tables we presented respectively the detailed results of the prediction of the OSA Toasty and Prune (condition A of optimization). The accuracy of the prediction was assessed by the r-squared values and highlighted a good prediction of the simple OSA Cut-grass ($r^2 = 0.68$, $p < 0.01$). However, no significant results were found for the other linear regressions. Several intensity values were not predicted by the model ($n < 16$). The absence of prediction was due to the absence of OQ required to elicit the perception of the given OSA. Indeed, the absence of odorants belonging to the OQ Lactonic in the wines from the grape variety Cabernet Franc led to a membership degree equals to 0 and thus resulted in an intensity predicted of the OSA Prune equal to 0 in the 8 Cabernet Franc wines. Focusing on the complex OSA, the condition A and B of optimization led to similar results, whereas condition C provide better prediction for the OSA Bell pepper and Cherry fresh.

In Figure 5, we presented the similarity between the evaluated and predicted value for the 7 simple OSA and the 8 complex OSA for the three conditions of optimization. In average, the similarity of the simple OSA is the highest (0.60). The ranking of the three conditions of optimization highlighted condition C as the most similar to the measured intensities (0.54) followed by condition A (0.40) and B (0.37).

We wondered if the predictive approach we developed could highlight the differences between the wines, i.e. separating the wines on the same qualitative attributes. To do so, we performed a principal component analysis (PCA) and hierarchical clustering on principle components (HCPC) on the OSA intensity evaluated and predicted for the 16 wines (Figure 6). Results from sensory evaluation showed a separation of the wines according to their grape varieties with the exception of CF3 grouped with the PN, and PN4 with the CF wines. CF wines were perceived as more Bell pepper, Blackcurrant, Cut-grass, Strawberry and Violet than the PN wines. Conditions A and B were also able to separate the wines according to their grape varieties, but with the exception of PN3 grouped with the CF. Similarity were also observed regarding the variables with CF wines perceived as more Blackcurrant and Strawberry but some variables were opposed compared to the results from sensory evaluation (Woody, Smoky). We further investigated the closeness of the PCA maps by calculating the RV coefficients (i.e. a multivariate generalization of the squared Pearson correlation coefficients) between the predicted values and the actual sensory ones according to the three conditions of otpimization. From this statistical test, the ranking of the three condition of optimization highlighted condition B as the closest to the evaluated profile (RV = 0.47),

followed by condition A (RV = 0.46) and lastly condition C (RV = 0.36). Indeed, condition C did only separate two PN wines from the other 14 wines.

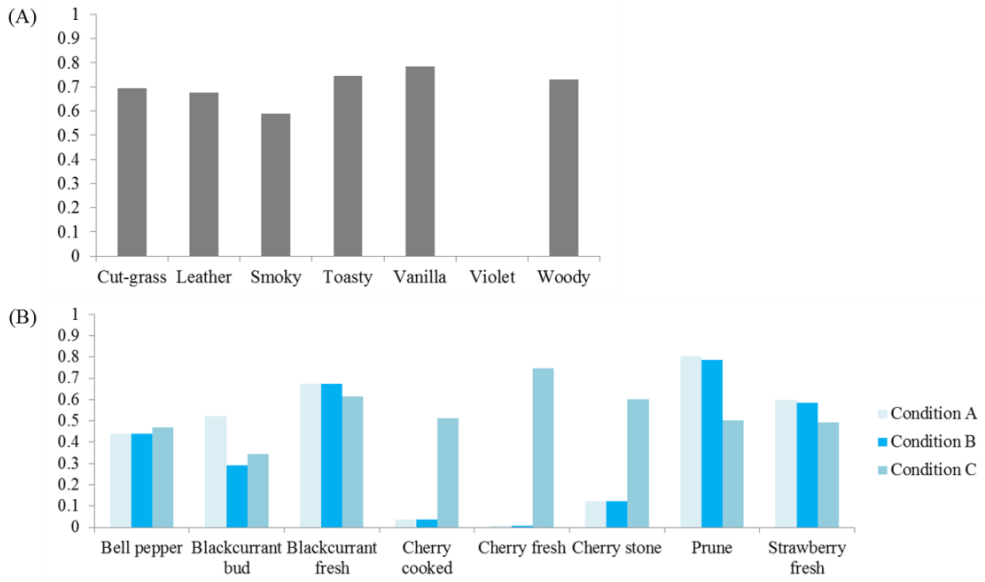
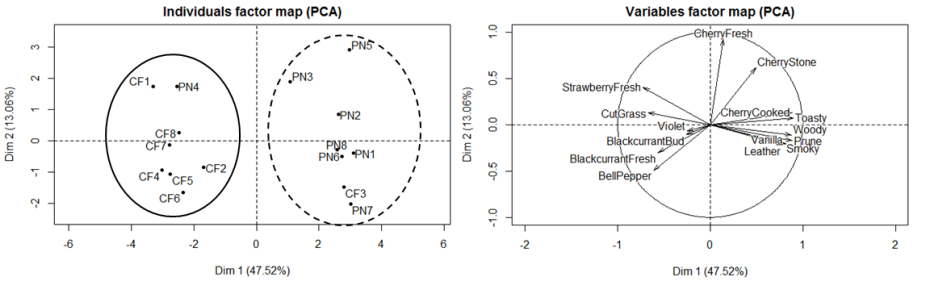
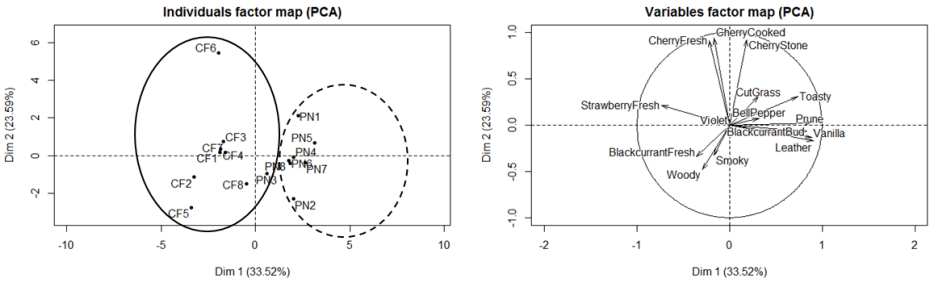


Figure 5: Similarities between the intensity evaluated by the sensory panel and intensity predicted by the model for the 15 OSA: (A) simple OSA and (B) complex OSA. Similarity varies from 0 to 1, 1 meaning that both intensities are equal. The three optimization conditions are shown (for condition C, the intensities predicted correspond to the means of the ten repetitions of the optimization).

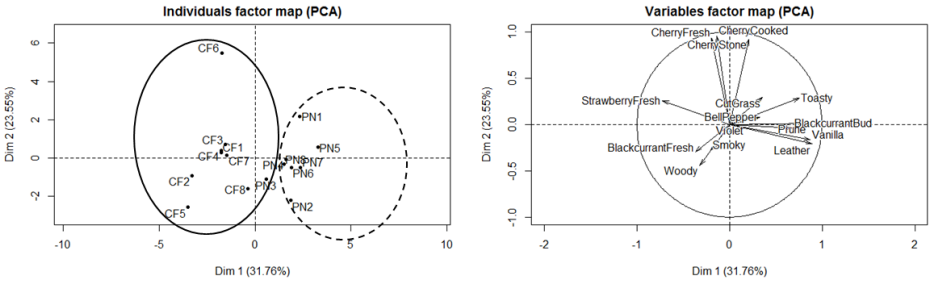
From sensory evaluation



From predictive approach (condition A)



From predictive approach (condition B)



From predictive approach (condition C)

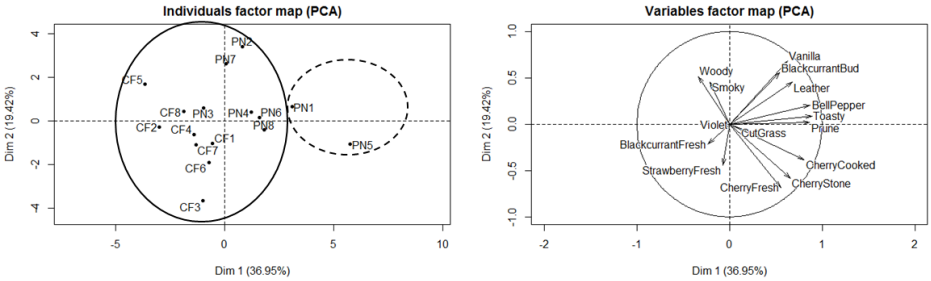


Figure 6: PCA maps based on the two first dimensions illustrating the configuration of the 16 wines (individuals) evaluated on 15 odor sensory attributes (variables): from sensory evaluation and from the predictive approach. Circles on the individuals factor map reflect the results of the HCPC. Means for each sensory descriptor are taken into account. PN: Pinot Noir wines, CF: Cabernet Franc wines.

Discussion

This research aimed at proposing an innovative strategy to predict the odor profile of food from its chemical composition. We chose to develop a holistic approach that combines knowledge and comprehension of scientific facts while relying upon computer science methods. This strategy especially took into consideration expertise to integrate non-linear perceptual computation of odorants' mixtures. To the best of our knowledge, the developed model constitutes the first attempt to predict quantitative odor description from molecular composition, but it is also the first report on knowledge-based artificial intelligence approach related to odor perception.

In contrast to most of previous modelling approaches in olfaction that concentrate on single odorants, the model was applied to a set of real food samples, namely wines. The model was able to predict the wines odor profiles through the estimation of the intensity of 15 odor sensory attributes (OSA). The previously developed Ontology for Odor Perceptual Space (OOPS; Roche et al., **Article 2 of this manuscript**) was used to establish the relationships between molecular content of food and perceptual concepts at the core of the expertise, while fuzzy logic was used to formalize expertise in order to make it useable for OSA prediction after optimization.

Regarding simple OSA, the similarity between the predicted and evaluated intensity was high (0.60). However correlations at the basis of simple OSA prediction were not significant, except for the OSA Cut-grass. Thus, results may highlight that the 6 remaining OSA identified as simple by the experts might be the results of combinations of OQ in the wine matrix. Indeed the OSA Leather can for instance be defined as complex with the OQ Animal and Smoky or Coumarinic.

Regarding complex OSA, they were predicted using fuzzy logic functions obtained following three conditions of optimization. Conditions varied according to the degree of expertise integrated in the model and thus on the definition of the fuzzy memberships functions. The membership functions for condition A relied on pure expertise: structure (the rules linking OSA to OQ) and parameters (the fuzzy function parameters) are fixed exclusively by experts. Condition B was a variant of condition A where the slopes of the fuzzy functions were optimized by a genetic algorithm. Results highlighted that conditions A and B were very similar because of the limited optimization possibilities. The membership functions for condition C relied also on expertise for the structure of the model (combination of OQ) but

not for the parameters. Those parameters were fully estimated using an evolutionary algorithm. Conditions A and B resulted in good similarities for some OSA (Blackcurrant fresh, Prune, Strawberry fresh) and a rather good classification of the wines (PCA, HCPC analyses). Condition C resulted in better similarity for some OSA (Cherry fresh, Cherry stone) and slightly better results in correlation between the intensity evaluated and predicted for three OSA (Bell pepper, Cherry fresh, and Prune; Figure S2). This finding tended to highlight that for some OSA like Cherry fresh, the expertise alone is not sufficient to predict what it perceived by the panelists. Hence, condition C might provide insights in combinations of OQ not explored by the experts. As illustrated in Figure 7, the optimized condition C highlighted that more OQ Fruity was needed to predict more accurately the intensity of the OSA Cherry fresh in the wines. Such result can be explained by the non-targeted approach we developed. Indeed, expertise was collected in a generic way and experts were not aware of the studied food matrix. However, the matrix of the product was shown to impact the perception of sensory attributes. Further investigations should be performed by presenting the results obtained to flavorists or wines experts in order to adapt the knowledge integrated in our modelling strategy. Such an approach might be of interest to underline which OSA are connected to the wine matrix and which one can be predicted independently. This could help to refine the experts rules related to odor associations that can underpin complex odor percepts.

The complementarity between machine learning and expertise allowed getting insights into olfactory attributes construction Indeed, the predictive strategy developed mimics the cognitive integration of odor information by reducing the information collected from odorants to the perception of OSA: a pool of fifty odorants was translated into 175 OD further linked to 20 OQ thanks to the OOPS, then, fuzzy logic was used to estimate the relationships between the OQ and the OSA.

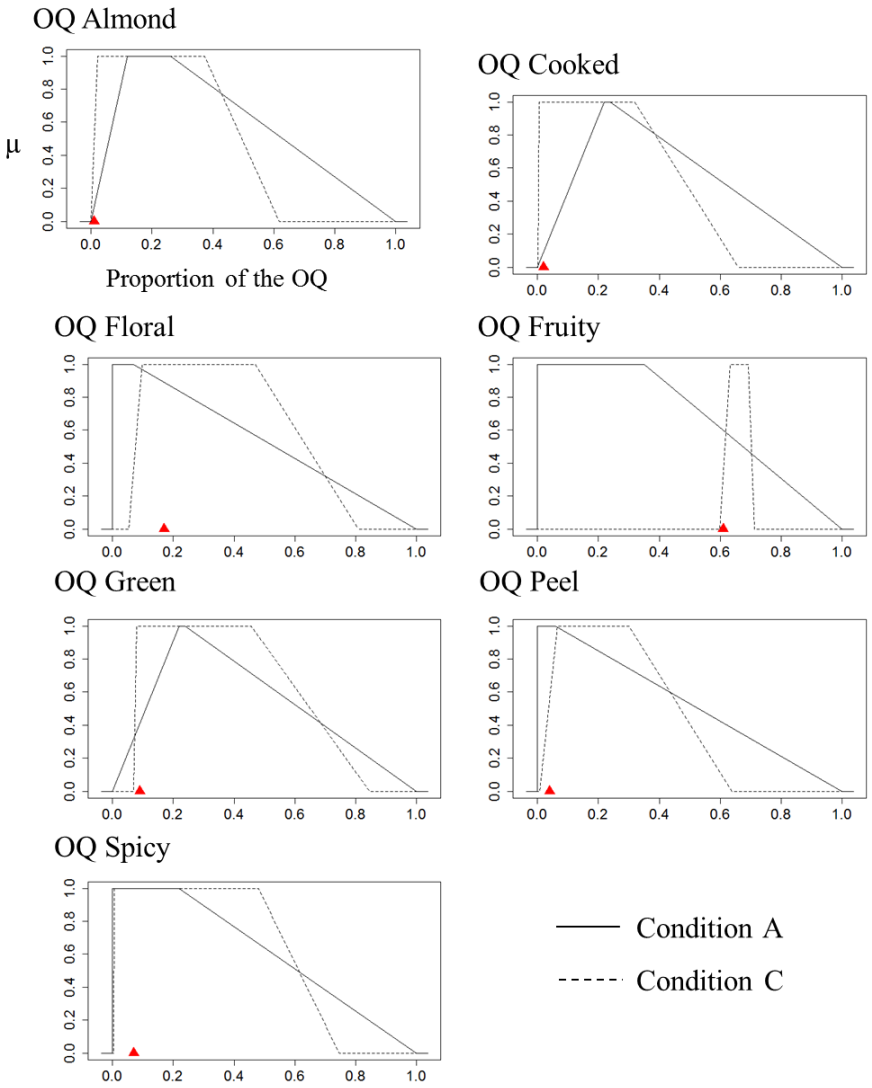


Figure 7: Comparison of the fuzzy membership functions of the OSA Cherry fresh between condition A and C.

Acknowledgments

This work was supported by grants from the Conseil Régional Bourgogne, Franche-Comté. Calculations were performed using HPC resources from PSIUN CCUB (Centre de Calcul de l'Université de Bourgogne). AR was supported by an INRA Thesis fellowship. We would like to thank Thomas Chabin for his help to compute on MATLAB.

References

Abraham, MH, Sánchez-Moreno, R, Cometto-Muñiz, JE & Cain, WS. 2012. An algorithm for 353 odor detection thresholds in humans. Chemical senses, 37(3):207-218. doi:10.1093/chemse/bjr094.

Arctander, S. 1969. Perfume and flavor chemicals, Volumes I and II. Montclair, NJ, USA.

Arlot, S & Celisse, A. 2010. A survey of cross-validation procedures for model selection. Statistics surveys, 4:40-79. doi:10.1214/09-SS054.

Bratolli, M, Cisternino, E, Dambruoso, PR, de Gennaro, G, Giungato, P, Mazzone, A, Palmisani, J & Tutino, M. 2013. gas chromatography analysis with olfactometric detection (GC-O) as a useful methodology for chemical characterization of odorous compounds. Sensors, 13:16759-16800. doi:10.3390/s131216759.

Cain, WS. 1969. Odor intensity: Differences in the exponent of the psychophysical function. Perception and psychophysics, 6(6):349-354. doi:10.3758/BF03212789.

Davidson, VJ, Ryks, J & Chu, T. 2001. Fuzzy models to predict consumer ratings for biscuits based on digital image features. IEEE transactions on fuzzy systems, 9(1):62-67. doi:10.1109/91.917115.

Debjani, C, Das, S & Das, H. 2013. Aggregation of sensory data using fuzzy logic for sensory quality evaluation of food. Journal of food science and technology, 50(6):1088-1096. doi:10.1007/s13197-011-0433-x.

Delahunty, CM, Eyres, G & Dufour, J-P. 2006. Gas chromatography-olfactometry. Journal of separation science, 29(14):2107-2125. doi:10.1002/jssc.200500509.

Dubois, D & Prade, H. 1980. Fuzzy sets and systems: Theory and applications. Academic press, New York, NY, USA.

Edwards, PA & Jurs, PC. 1989. Correlation of odor intensities with structural properties of odorants. Chemical senses, 14(2):281-291. doi:10.1093/chemse/14.2.281.

Hornung, DE & Enns, MP. 1986. The contributions of smell and taste to overall intensity: A model. Perception and psychophysics, 39(6):385-391.doi:10.3758/BF03207066.

Hudson, R. 2000. Odor and odorant: A terminological clarification. Chemical senses, 25(6):693. doi:10.1093/chemse/25.6.693.

Escudero, A, Gogorza, B, Melús, MA, Ortín, N, Cacho, J & Ferreira, V. 2004. characterization of the aroma of a wine from maccabeo. Key role played by compounds with low odor activity values. Journal of agricultural and food chemistry, 52(11):3516-3524. doi:10.1021/jf035341l.

Ferrari, G, Lablanque, O, Cantagrel, R, Ledauphin, J, Payot, T, Fournier, N & Guichard, E. 2004. Determination of key odorant compounds in freshly distilled cognac using GC-O, GC-MS, and sensory evaluation. Journal for agricultural and food chemistry, 52(18):5670-5676. doi:10.1021/jf049512d.

Ferreira, V, Ortín, N, Escudero, A, López, R & Cacho, J. 2002. Chemical characterization of the aroma of Grenache rosé wines: Aroma extract dilution analysis, quantitative determination, and sensory reconstitution studies. Journal of agricultural and food chemistry, 50(14):4048-4054. doi:10.1021/jf0115645.

Ferreira, V, Sáenz-Navajas, MP, Campo, E, Herrero, P, de la Fuente, A & Fernández-Zurbano, P. 2016. Sensory interactions between six common aroma vectors explain four main red wine aroma nuances. Food chemistry, 199:447-456. doi:10.1016/j.foodchem.2015.12.048.

Folorunso, O, Ajayi, Y & Shittu, T. 2009. Fuzzy-Rule-Based approach for modeling sensory acceptabitity of food products. Data science journal, 8:70-77. doi:10.2481/dsj.007-006.

Grosch, W. 2001. Evaluation of the key odorants of foods by dilution experiments, aroma models and omission. Chemical senses, 26(5):533-545. doi:10.1093/chemse/26.5.533.

Guang, W, Baraldo, M & Furlanut, M. 1995. Calculating percentage prediction error: A user's note. Pharmacological research, 32(4):241-248. doi:10.1016/S1043-6618(05)80029-5.

Guth, H. 1997. Quantitation and sensory studies of character impact odorants of different white wine varieties. Journal of agricultural and food chemistry, 45(8): 3027-3032. doi:10.1021/jf970280a.

Hansen, N. 2006. The CMA evolution strategy: A comparing review. In: Towards a New Evolutionary Computation Studies in Fuzziness and Soft Computing, Volume 192, eds Lozano, JA, Larrañaga, P, Inza, I, Bengoetxea E. Springer, Berlin, Heidelberg. 75-102.

Ioannou, I, Perrot, N, Hossenlopp, J, Mauris, G & Trystram, G. 2002. The fuzzy set theory: A helpful tool for the estimation of sensory properties of crusting sausage appearance by a single expert. Food quality and preference, 13:589-595. doi:10.1016/S0950-3293(02)00045-9.

Jaubert, J-N, Tapiero, C & Dore, J-C. 1995. The field of odors: Toward a universal language for odor relationships. Perfumer and flavorist, 20:1-16.

Keller, A, Gerkin, RC, Guan, Y, Dhurandhar, A, Turu, G, Szalai, B, Mainland, JD, Ihara, Y, Yu, CV, Wolfinger, R, Vens, C, Schietgat, L, De Grave, K, Norel, R, DREAM Olfaction Prediction Consortium, Stolovitzky, G, Cecchi, GA, Vosshall, LB & Meyer, P. 2017. Predicting human olfactory perception from chemical features of odor molecules. Science, 355(6327):820-826. doi:10.1126/science.aal2014.

Khan, RM, Luk, C-H, Flinker, A, Aggarwal, A, Lapid, H, Haddad, R & Sobel, N. 2007. Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. Journal of neuroscience, 27(37):10015-10023. doi:10.1523/JNEUROSCI.1158-07.2007.

Kumar, R, Kaur, R, Auffarth, B & Bhondekar, AP. 2015. Understanding the odour spaces: A step towards solving olfactory stimulus-percept problem. Plos one, 10(10): e0141263. doi:10.1371/journal.pone.0141263.

Lee, S-J, & Noble, AC. 2003. Characterization of odor-active compounds in Californian Chardonnay wines using GC-olfactometry and GC-mass spectrometry. Journal of agricultural and food chemistry, 51(27):8036-8044. doi:10.1021/jf034747v.

Loison, A, Symoneaux, R, Deneulin, P, Thomas-Danguin, T, Fant, C, Guérin, L & Le Fur, Y. 2015. Exemplarity measurement and estimation of the level of interjudge agreement for two categories of French red wines. Food quality and preference, 40:240–251. doi:10.1016/j.foodqual.2014.10.001.

Lorrain, B, Ballester, J, Thomas-Danguin, T, Blanquet, J, Meunier, JM & Le Fur, Y. 2006. Selection of potential impact odorants and sensory validation of their importance in typical Chardonnay wines. Journal of agricultural and food chemistry, 54(11):3973-3981. doi:10.1021/jf052350g.

Luebke, W. 1980. The good scents company. Available at: http://www.thegoodscentscompany.com/index.html

Mauris, G, Benoit, E & Foulloy, L. Fuzzy symbolic sensors – From concept to applications. Measurement, 12(4):357-384. doi:10.1016/0263-2241(94)90067-1.

Mitchell, M. 1996. An introduction to genetic algorithms. MIT press, Cambridge, UK.

Mukhopadhyay, S, Majumdar, GC, Goswami, TK & Mishra, HN. 2013. Fuzzy logic (similarity analysis) approach for sensory evaluation of chhana podo. LWT - Food science and technology, 53(1):204-210. doi:10.1016/j.lwt.2013.01.013.

Pérez-Silva, Odoux, E, Brat, P, Ribeyre, F, Rodriguez-Jimenes, G, Robles-Olvera, V, García-Alvarado, MA & Günata, Z. 2006. GC-MS and GC-olfactometry analysis of aroma compounds in a representative organic aroma extract from cured vanilla (Vanilla planifolia G. Jackson) beans. Food chemistry, 99(4):728-735. doi:10.1016/j.foodchem.2005.08.050.

Perrot, N, Ioannou, I, Allais, I, Curt, C, Hossenlopp, J & Trystram, G. 2006. Fuzzy concepts applied to food product quality control: A review. Fuzzy sets and systems, 157(9):1145-1154. doi:10.1016/j.fss.2005.12.013.

Perrot, N, Trelea, IC, Baudrit, C, Trystram, G & Bourgine, P. 2011. Modelling and analysis of complex food systems: State of the art and new trends. Trends in food science and technology, 22(6):304-314. doi:10.1016/j.tifs.2011.03.008.

Pollien, P, Ott, A, Montigon, F, Baumgartner, M, Muñoz-Box, R & Chaintreau, A. 1997. Hyphenated headspace-gas chromatography-sniffing technique: Screening of impact odorants and quantitative aromagram comparisons. Journal of agricultural and food chemistry, 45:2630-2637. doi:10.1021/jf960885r.

Shang, L, Liu, C, Tomiura, Y & Hayashi, K. 2017. Machine-learning-based olfactometer: Prediction of odor perception from physicochemical features of odorant molecules. Analytical chemistry, 89(22):11999-12005. doi:10.1021/acs.analchem.7b02389.

Shinde, KJ & Pardeshi, IL. 2014. Fuzzy logic model for sensory evaluation of commercially available jam samples. Journal of ready to eat food, 1(2):78-84.

Sicard, M, Baudrit, C, Leclerc-Perlat, M-N, Wuillemin & Perrot, N. 2011. Expert knowledge integration to model complex food processes. Application on the camembert cheese ripening process. Experts systems with applications, 38(9):11804-11812. doi:10.1016/j.eswa.2011.03.068.

Snitz, K, Yablonka, A, Weiss, T, Frumin, I, Khan, RM & Sobel, N. 2013. Predicting odor perceptual similarity from odor structure. Plos computational biology, 9(9):e1003184. doi:10.1371/journal.pcbi.1003184.

Stone, H & Sidel, JL. 2004. Sensory evaluation practices: Third edition, ed Taylor, SL. Elsevier academic press.

Tan, J, Gao, X & Gerrard, DE. 1999. Application of fuzzy sets and neural networks in sensory analysis. Journal of sensory studies, 14:119-138. doi:10.1111/j.1745-459X.1999.tb00108.x.

Thomas-Danguin, T, Sinding, C, Romagny, S, El Mountassir, F, Atanasova, B, Le Berre, E, Le Bon, A-M & Coureaud, G. 2014. The perception of odor objects in everyday life: A review on the processing of odor mixtures. Frontiers in psychology, 5:504. doi:10.3389/fpsyg.2014.00504.

Veinand, B. 2015. Sensory testing with flavourists: Challenges and solutions. In: Rapid sensory profiling techniques and related methods: Applications in new product development and consumer research, eds Delarue, J, Lawlor, JB, Rogeaux, M. Woodhead publishing. doi:10.1533/9781782422587.3.383. 383-399.

Zadeh, LA. 1965. Fuzzy sets. Information and control, 8(3):338-353. doi:10.1016/S0019-9958(65)90241-X.

Zadeh, LA. 1975. Fuzzy logic and approximate reasoning. Synthese, 30:407-428. doi:10.1007/BF00485052.

Zheng, Y, Sun, B, Zhao, M, Zheng, F, Huang, M, Sun, J, Sun, X & Li, H. 2016. Characterization of the key odorants in Chinese zhima aroma-type baijiu by gas chromatography-olfactometry, quantitative measurements, aroma recombination, and omission studies. Journal of agricultural and food chemistry, 64(26):5367-5374. doi:10.1021/acs.jafc.6b01390.

Supplementary data

Figure S1: Linear regression between the intensities evaluated and predicted for the 7 simple OSA. The r-squared value (r^2), p-value (p) and number of observations (n, excluding NA) of the regression are presented. The plot of the OSA Violet is not shown because the model did not predict the intensity of this OSA (n = 0).

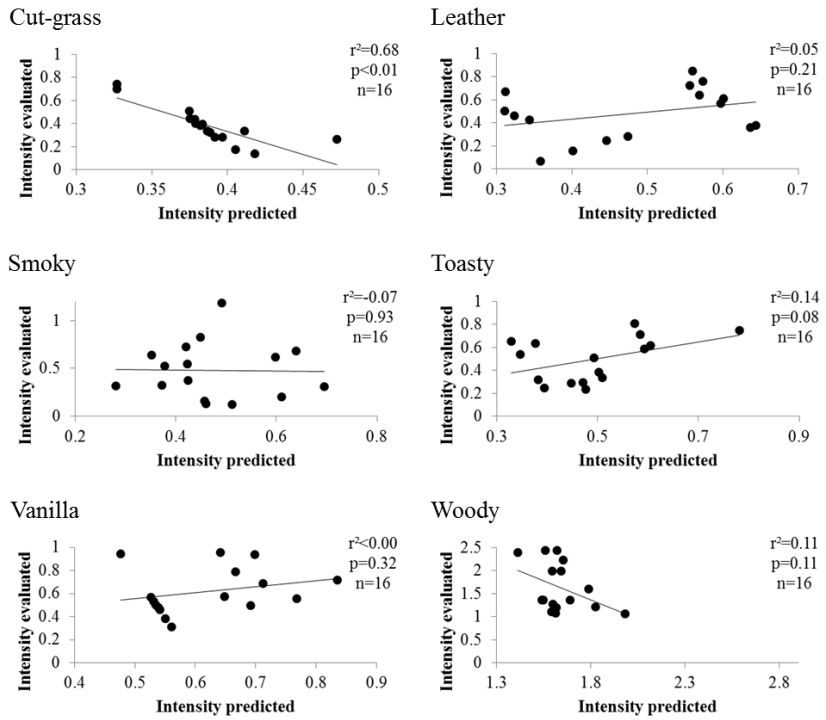


Figure S2: Linear regression between the intensities evaluated and predicted for the 8 complex OSA. The r-squared value (r^2), p-value (p) and number of observations (n, excluding NA) of the regression are presented.

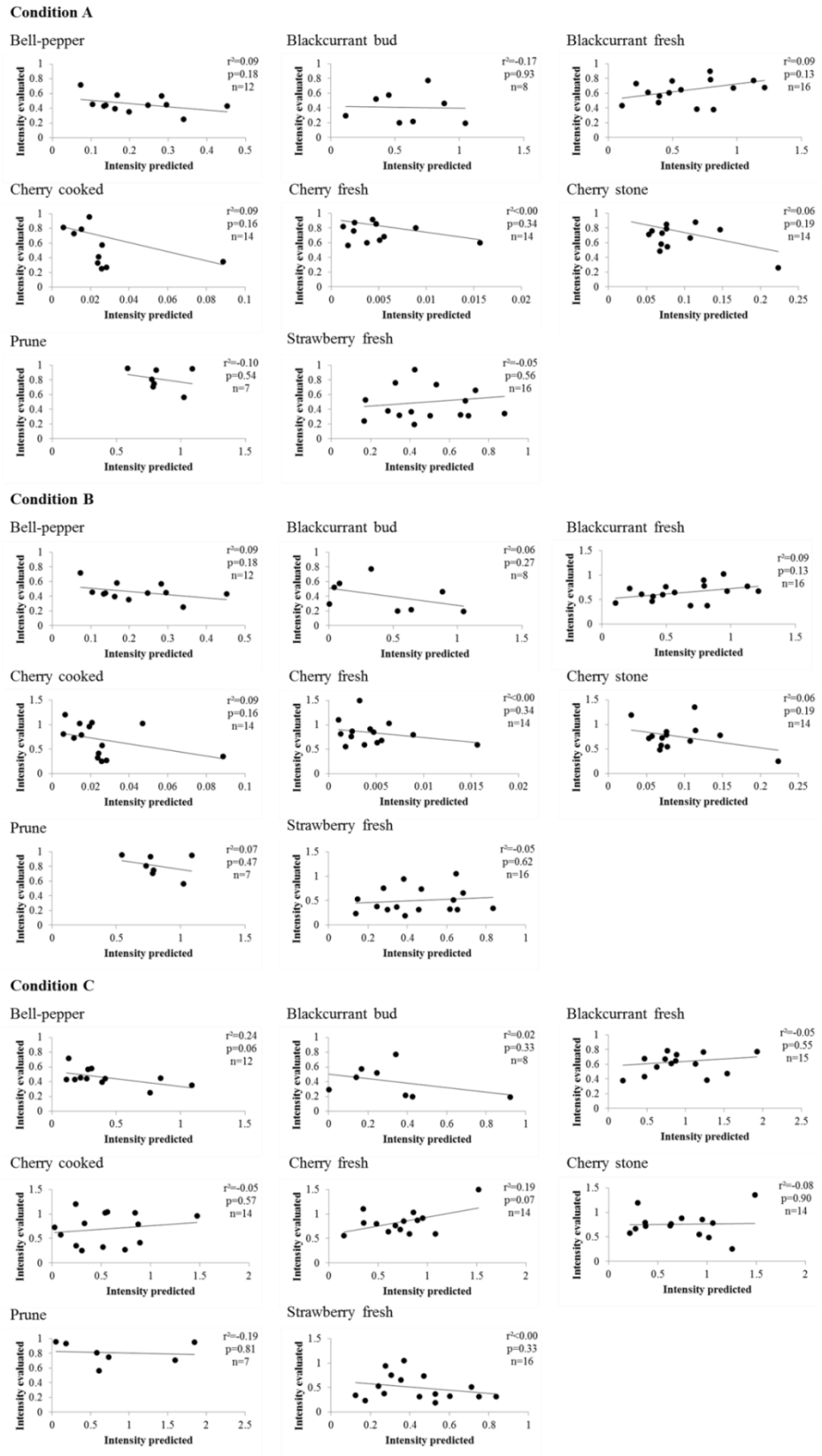


Table S1: Results of the linear regression between the intensity evaluated of the 7 simple OSA and the proportion of the corresponding OQ for 15 wines. The 15 wines correspond to the leave-one-out cross-validation (LOOCV) to predict the OSA intensity of the wine PN1. The r-squared value (r^2) and p-value (p) of the regression are presented.

	Linear regression between intensity evaluated and OQ proportion
Cut-grass	$r^2 = -0.04$, $p = 0.53$
Leather	$r^2 = 0.22$, $p = 0.05$
Smoky	$r^2 = 0.07$, $p = 0.18$
Toasty	$r^2 = 0.27$, $p = 0.03$
Vanilla	$r^2 = 0.09$, $p = 0.14$
Violet	$r^2 = 0$, $p = NA$
Woody	$r^2 = -0.07$, $p = 0.74$

Table S2: Fuzzy parameters a1 to a4 of the fuzzy membership functions linking the proportion of the five OQ Almond, Cooked, Floral, Fruity, Green, Peel, Spicy to the intensity of the OSA Prune according to the three conditions of optimization. For conditions 3 and 4, the fuzzy parameters are the results of the first repetition of the optimization.

Condition A	a1	a2	a3	a4
Almond	0	0	0.33	1
Cooked	0	0	0.75	1
Fruity	0	0	0.25	1
Honey	0	0	0.06	1
Lactonic	0	0	0.67	1
Condition B (Rep1)	a1	a2	a3	a4
Almond	0	0	0.33	0.61
Cooked	0	0	0.75	1
Fruity	0	0	0.25	1
Honey	0	0	0.06	0.97
Lactonic	0	0	0.67	1
Condition C (Rep1)	a1	a2	a3	a4
Almond	0	0	0.43	0.68
Cooked	0	0.03	0.32	0.72
Fruity	0.85	0.94	0.97	0.99
Honey	0.02	0.04	0.32	0.60
Lactonic	0	0.03	0.37	0.66

Table S3: Intensity predicted for the OSA Toasty for the 16 wines. PN: Pinot Noir wines, CF: Cabernet Franc wines. The maximal score for the intensity is 10. The percentage prediction error is calculated for each wine between the intensity evaluated by the panel and the one predicted for the OSA Toasty. The similarity (sim) between the intensities evaluated and predicted for the 16 wines is calculated following the Ruzicka similarity (Eq 6). Finally we perform a linear regression between the sensory and predicted OSA intensities and determine the r-squared value (r^2) and p-value (p) of the regression.

	Intensity evaluated	Intensity predicted	Percentage prediction error
PN1	0.61	0.60	1.61
PN2	0.51	0.49	3.15
PN3	0.65	0.33	49.56
PN4	0.38	0.50	31.80
PN5	0.75	0.78	4.23
PN6	0.59	0.59	0.88
PN7	0.71	0.58	18.24
PN8	0.81	0.57	29.17
CF1	0.29	0.47	61.48
CF2	0.54	0.35	35.70
CF3	0.63	0.38	40.62
CF4	0.23	0.48	103.90
CF5	0.32	0.38	19.92
CF6	0.34	0.51	50.72
CF7	0.25	0.39	60.24
CF8	0.29	0.45	56.73
	sim = 0.75, r^2 = 0.14, p = 0.08		

Table S4: Intensity predicted for the OSA Prune according to the condition A of optimization for the 16 wines. PN: Pinot Noir wines, CF: Cabernet Franc wines. The maximal score for the intensity is 10. The percentage prediction error is calculated for each wine between the intensity evaluated by the panel and the one predicted for the OSA Prune. *NA* means that the percentage prediction error calculation is not applicable because the intensity of the OSA was not predicted by the model. The similarity (sim) between the intensities evaluated and predicted for the 16 wines is calculated following the Ruzicka similarity (Eq 6). Finally we perform a linear regression between the evaluated and predicted OSA intensities and determine the r-squared value (r^2) and p-value (p) of the regression.

	Intensity evaluated	Intensity predicted	Percentage prediction error
PN1	0.95	1.09	14.7
PN2	0.94	0.81	13.8
PN3	0.74	*Not Predicted*	*NA*
PN4	0.56	1.02	8.1
PN5	0.71	0.78	9.9
PN6	0.81	0.77	4.9
PN7	0.95	0.58	38.9
PN8	0.75	0.79	5.3
CF1	0.13	*Not Predicted*	*NA*
CF2	0.63	*Not Predicted*	*NA*
CF3	0.89	*Not Predicted*	*NA*
CF4	0.38	*Not Predicted*	*NA*
CF5	0.39	*Not Predicted*	*NA*
CF6	0.49	*Not Predicted*	*NA*
CF7	0.55	*Not Predicted*	*NA*
CF8	0.55	*Not Predicted*	*NA*
	sim = 0.80, r^2 = 0.27, p = 0.02		

III. Partial discussion

Within the **ARTICLE 3**, we proposed a machine learning approach relying on flavorists' expertise to predict the odor profile of wines on the basis of their chemical composition. The results were promising and supported the idea that converting the chemical composition of food products into their corresponding odor description (odor descriptors, OD and odor qualities OQ) and using expert knowledge to rebuild overall odor sensory attributes (OSA) is a good modelling strategy.

In Figure 5.1, the estimated odor profiles of the two wines PN8 and CF7 (respectively identified as PN-A and CF-A in Roche et al., **Article 2 of this manuscript**) obtained with the integrative approach based on fuzzy logic (condition B of optimization) are compared with those obtained from the experimental sensory characterization of the wines (Villière et al., **Data paper included in this manuscript**). The comparison between the predicted and experimental profiles showed a good agreement between both profiles except for the Cherry-related attributes (Cherry cooked, Cherry fresh, and Cherry stone) for both wines. Moreover, the OSA Violet could not be predicted for both wines and the OSA Bell pepper, Blackcurrant bud, and Prune for the CF-A wine (set arbitrarily at the minimum value in Figure 5.1). For the OSA Violet, no identified molecules in the wines carried the OD that are part of the OQ Violet, which is directly linked to the OSA Violet since Violet is a simple OSA according to expert knowledge. For the OSA Bell pepper and Prune, no identified molecules in the CF-A wine carried the OD that are part of the OQ Lactonic, which was used by the fuzzy rules for these two OSA; for the OSA Blackcurrant bud, no molecules carried the OD part of the OQ Vanilla used in the Blackcurrant bud fuzzy rule.

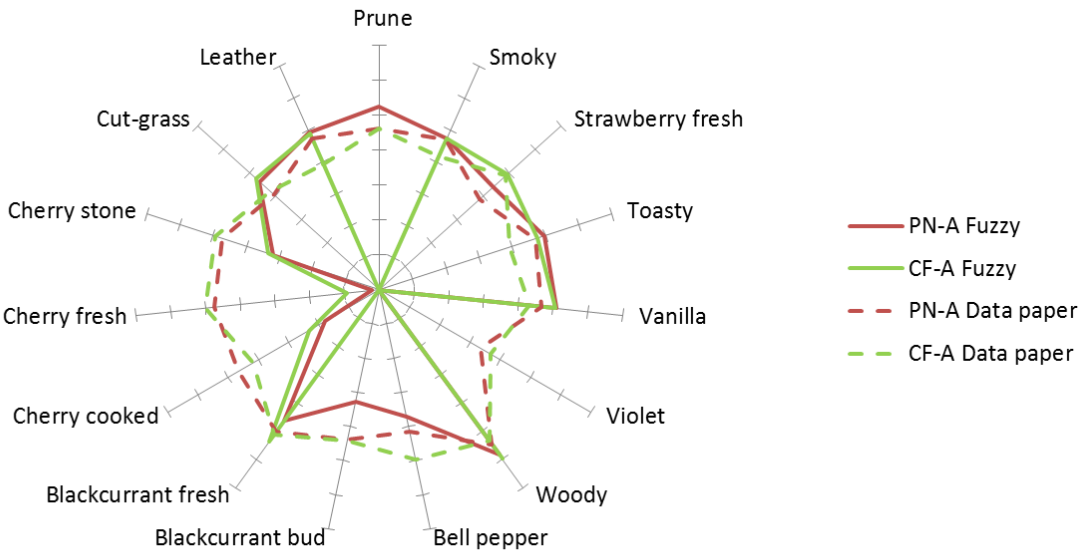


Figure 5.1: Radar plots illustrating the profile of the OSA in the PN-A and CF-A wines. Proportions of OSA are plotted in log scales (arbitrary units). Results of the fuzzy optimization were obtained with the condition B. PN-A corresponds to the wine PN8 and CF-A to the wine CF7.

In conclusion, even if the model might be improved in order to more accurately reflect the sensory aspects of wines; it was proven able to predict the intensity of several wines' OSA on the basis of their chemical composition. Several perspectives can be proposed to improve the accuracy of the model. On the one hand, various machine learning parameters can be modified (optimization condition, fuzzy rules parameters). On the other hand, the increase of the dataset of food products' chemical composition and sensory evaluation may be considered. Nevertheless, we have to keep in mind that such data might contain a part of uncertainty related to incomplete volatile compounds identification or inter-individual variability among sensory panelists.

This innovative work opens up the scope of possibilities to integrate expertise along with analytical data to predict the sensory outcomes of complex odor sources. The following general discussion aims to point the limitations of the approach compared to the previous studies and to highlight several perspectives.

GENERAL DISCUSSION & PERSPECTIVES

The work presented in this PhD thesis manuscript aimed to improve food flavor analysis by developing efficient approaches to predict the odor resulting of complex odorants' mixtures such as those elicited by wines odors. To do so, we worked on two distinct but complementary approaches which represent the two axes of the manuscript (Figure D.1).

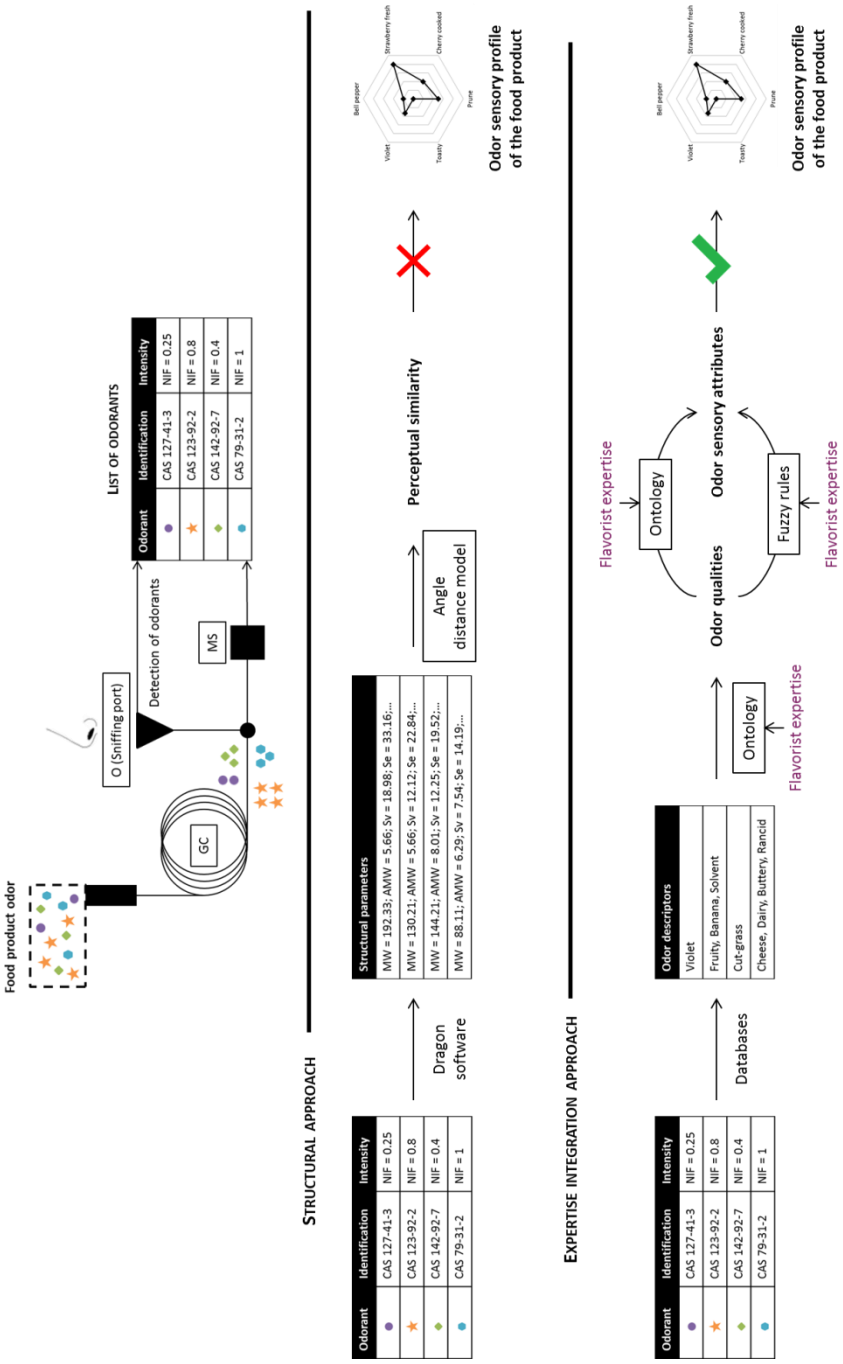


Figure D.1: The two approaches presented in this manuscript in order to improve food flavor analysis.

The first axis of the manuscript explored a structural approach to predict similarity between odor stimuli. We tested and upgraded the angle distance model developed by Snitz et al. (2013) on various odorants mixtures (up to 6 odorants). Results highlighted that taking into account the ratio of odorants within the mixtures improved the prediction accuracy of the model and furthermore allowed to account for the elemental and configural processing of odorants' mixtures by the olfactory system (Roche et al., **Article 1 of this manuscript**). The application of the newly developed ratio-weighted angle distance model to predict the similarity between highly complex odorants mixtures, namely wines, resulted in a significant prediction ability highlighted by the significant correlation between wines odor similarity and angle distances obtained with the ratio-weighted angle distance model. However, the level of prediction remained weak (low correlation value) in part because of the poor wines odor similarity data reliability, which had to be inferred from experimental sensory profiles of the wines. Altogether, this axis focused on predicting the similarity among mixtures of odorants estimated on the basis of the molecular structure of odorants, which is a first, but major step in getting insights in the prediction of the perceptual outcome of odorants' mixtures. However, the aim of this thesis was to go beyond the similarity dimension and to predict what an odorants' mixture will smell, that is to provide verbal descriptors with a quantitative aspect as in an odor profile, usually obtained after quantitative descriptive analysis.

The second axis of the manuscript explored a modelling approach integrating flavor expertise along with chemical data to predict the odor quality of complex (real) odorants' mixtures. The expertise collection from expert flavorists led to the construction of the Ontology for Odor Perceptual Space (OOPS), detailed in Roche et al. (**Article 2 of this manuscript**). The OOPS aimed to represent the relationships between three levels of the odor perceptual space: odor descriptors (OD) used to describe single odorant and available in several databases, odor qualities (OQ) as the corpus used by flavorists to construct their operational perceptual space, and odor sensory attributes (OSA) used by sensory trained panels to describe food products. The ontology was used as a predictive tool to estimate wines' odor profiles (i.e. OSA proportions) on the basis of their chemical composition. Despite interesting prediction abilities, which need to be further confirmed, this approach was limited since it relied more on a classification into odor categories and considered the relationships between each level of the odor perceptual space as additive, which do not reflect the behavior of the system and the actual integration processes of complex odorants' mixtures stimuli.

Thus, in order to take into account non-additive relationships, we integrated the OOPS in a broader model relying on fuzzy logic and optimization (Roche et al., **Article 3 of this manuscript**). The OOPS allowed the translation of chemical stimuli into OQ proportions. Then OQ proportions were modeled according to the expertise of flavorists, who identified rules in odor mixtures perception to elicit specific OSA. These rules were formalized using fuzzy logic and optimized owing to data-driven optimization procedures. The fuzzy modelling was proven efficient to predict the odor profiles of 16 red wines; the input of the model being the chemical composition of the wines and the output being the intensity of several OSA.

In the following paragraphs, a general discussion of the different approaches considered in this thesis work, their limitations, and corresponding perspectives are proposed.

I. Chemical and sensory data

The food odor data used throughout the manuscript to test the modelling approaches were obtained from the chemical and sensory characterization of 16 red wines.

The chemical analysis of the wines presents some gaps. Indeed, the results of the odorants' identification listed 49 odorants and 34 non-identified compounds (Villière et al., 2018). The compounds non-identified might have impacted the odor perception of the wines, but it was not possible to assign them an odor description from the databases. Therefore, they were not considered during the modelling studies. One solution to overcome this issue would have been to take into account the odor description collected from the subjects during the gas chromatography - olfactometry (GC-O) step. However, the descriptors list provided during GC-O was not exhaustive and high inter-individual variability was observed among the subjects; a first mandatory step would then be to align the GC-O vocabulary with the OD found in databases.

The sensory evaluation of the wines present also some bias. Even if the panelists who evaluated the odors of the 16 red wines were highly trained for wine profiling, their use of the rating scales was not highly consensual and thus resulted in a high inter-individual variability. In addition, several attributes (i.e. OSA) were found not discriminant between the wines although they were selected for this aim (Villière et al., **Data paper included in this**

manuscript). These findings may reflect either low differences between the wines, which would have increased the prediction difficulty, or a lack of consensus with regards to many attributes. One possibility to address this last issue would be to pre-process the sensory data, using for instance the mixed assessor model (MAM) proposed by Brockhoff et al. (2015) to produce normalized OSA intensity values, which may reduce uncertainty propagation within the models.

II. Structural approach to create odor profiles

The structural approach we developed aimed to predict the perceptual similarity among various odorants' mixtures, especially those released by real food products.

The estimation of similarity is an interesting approach for odor quality prediction since it avoids the use of semantic descriptors (Snitz et al., 2013). Indeed, semantic descriptors are often considered subjective and may be prone to several biases as illustrated in this manuscript by the difficulty in odor naming and the rather low consensus regarding the databases providing the odor description of molecules. Hence, the results of similarity approach are relative by comparing two stimuli in a discriminative process, which is less prone to language-related biases (Wise et al., 2000). In this thesis work, the angle distance model was found rather suitable for predicting the odor similarity among complex stimuli knowing their molecular structure.

Therefore, one possibility to use such predictive approach within the aroma analysis path would be to estimate similarity or dissimilarity in simulated omission tests, which would help identifying key odorants. Indeed, sub-mixtures tested within the angle distance model in Chapter 3 may be seen as omission situations and the model was found fully suitable for such situations. In the aroma analysis procedure, in order to assess the odor impact of odorants embedded in a food product, stimuli made of odorants alone or in mixtures are experimentally created and compared to a reference in sensory evaluation. This is a long process and most of the time all the possible combinations cannot be tested. For example, Xiao et al. (2017) identified 36 volatiles in mandarin juices and they tested only 23 omission models to verify the contribution of certain volatiles. However, if one wanted to test every possible omission models, the sensory evaluation of thousands models would be required. Thus, the angle

distance approach could be applied as a quick simulation approach to estimate the impact of each component or components' mixtures on the overall perception of a food product. In that case the predicted similarity between the reduced mixtures stimulus and the complete formulation would help to determine the most potent odorants or odorants' combinations (Romagny et al., 2018). If the similarity between the stimuli decreases, the omitted odorants might impact the overall odor of the food product. Further work would be needed to verify the validity of this hypothesis.

With the aim to predict the odor of food product, semantic information remains the most suitable one because we have previously mentioned that odors are commonly qualified by terms related to odor sources. In spite of previous issues evoked concerning verbal descriptions of odors, quantitative descriptive analysis of odor, relying on the quantitative rating of odor sensory attributes (OSA) was proven to be a reliable method to characterize a food odor with verbal terms (Strigler et al., 2009). Thus, for further work we thought about using the structural approach to predict the semantic profile of molecules. To do so, we propose to rely on data from the Atlas of odor character profiles (Dravnieks, 1985) which contains the sensory description of 144 molecules. Each molecule was rated for 146 semantic descriptors (e.g. Animal, Cinnamon, Fruity, Meaty, Metallic, Smoky) on a 0-5 scale. The objective will be to determine the structural vector best characterizing each one of the 146 semantic descriptors.

To do so, optimization techniques could be used to maximize the correlation between (i) the angle distance between a given molecule structural vector with the semantic descriptor structural vector and (ii) the semantic score of this molecule provided in the Atlas of odor character profiles, which would result in 144 values corresponding to the 144 molecules. Once the 146 semantic descriptor structural vectors obtained, it would become possible to calculate, for any given stimulus (single odorant or odorants' mixture), the angle distance between the structural vector of this stimulus and each one of the 146 semantic structural vectors, which results in 146 values of similarity that represent the odor profile of the stimulus. The result of the structural approach will then become a semantic profile (using the descriptors provided in the Atlas) instead of a similarity between stimuli. This proposed perspective may allow predicting semantic odor profile of molecules or mixtures on the basis of their molecular structure.

III. Expertise integration approach

The construction of the OOPS and the integrative approach based on fuzzy logic relied on flavorists' expertise. The following discussion is centered on expertise integration and the comparison between both approaches of integration.

1. Data collection

The expertise was collected in order to establish the relationships between the three dimensions of the odor perceptual space (OD, OQ, and OSA). To do so, the starting point of the expertise elicitation was based on 15 OSA related to the sensory characterization of red wines. Therefore, the number of OSA considered did not cover the whole olfactory space. Still, there is no consensus about the minimum number of descriptors or odor categories that are necessary to represent the olfactory space. For instance, Dravnieks (1985) proposed the use of 146 semantic descriptors to fully describe a set of 144 molecules. In a recent study, Kumar et al. (2015) studied several databases of odorants' description and developed a graphical method to find the similarity between perceptual descriptors. They found 7 major communities each representing a certain class of perceptual quality (Herb-Wood, Green-Vegetable, Floral, Citrus-Oil, Earth-Meat and Sweet-Fruit). However, within each major category, several sub-categories are likely to be required to cover the olfactory space with a sufficient accuracy. It would be interesting to select several of these categories as OSA to complete the expertise included in the ontology. Indeed, these OSA could be proposed to the experts to collect data on the relationships between these new attributes and the existing OQ or to define new ones if necessary. This is a major advantage of the ontology formalism since it is flexible and thus the tool can be adapted to different food products and/or extended to embed new knowledge.

2. Genericity

The expertise embedded in our approach was collected toward a non-targeted food product. Indeed during the expertise collection, the experts were not informed that the OSA they had to

consider were related to wine. Thus, the collected expertise can be viewed as generic and can be used for the study of various food products.

During the elicitation process, the mental concept evoked for each expert for a given OSA might have been different. For example, the concept Strawberry fresh might have been mentally processed within the framework of different matrices such as yoghurts or candies, for instance according to the flavorist's specific flavor application specialty. To evaluate the importance of such an issue, a perspective would be to compare the generic expertise against the non-generic one. As mentioned in the Chapter 2, the expertise was also collected from the expert flavorists after they had been informed that the OSA were related to a wine matrix (the data were not used in the thesis). Thus, comparing results between both types of data would be of interest to assess the genericity of the expertise integration approach. Furthermore, one can expect that the knowledge gathered toward a non-targeted or targeted product might highlight different rules of odor mixtures perception.

Another issue related to the food matrix is linked to the fact that the composition of a food product impacts the release of the odorants, thus the proportion of the odorants reaching the olfactory mucosa, and therefore the perception of the food odor (Coureaud et al., 2011; Déléris et al., 2012). The release kinetics of odorants depends on the matrix properties (structure, viscosity) and its constituents. For example, interactions between odorants and fats (van Ruth et al., 2002) or proteins (Tromelin et al., 2006) were pointed out in the literature. Thus, non-volatiles compounds can indirectly impact odor perception of food products. Concerning wines and spirits, ethanol was found to have an impact on the flavor profile of the products (Le Berre et al., 2007; Ickes & Cadwallader, 2017). For instance, decreasing the ethanol concentration in wines led to an increase of the fruity and flowery odors (Grosch, 2001). As soon as expert flavorists usually create flavors for a given application, it is likely that their expertise can also empirically take into account these effects so that it cannot be excluded that the collection of expertise out of a food product context may face some limitations.

3. Experts' validation of the fuzzy formalism

The integrative approach based on fuzzy logic was built on flavorists' expertise. As we highlighted in the methodology, several choices have to be made when constructing fuzzy

models: data and model structuration, choice of operators, parameter ranges, fuzzy rules, defuzzification method… Human choices have thus a major impact at various levels of the fuzzy formalism (Perrot & Baudrit, 2013).

In order to validate the fuzzy model presented in Roche et al. (**Article 3 of this manuscript**), one perspective would be to validate the model created with the experts in an interactive manner. By showing them the fuzzy rules and results we obtained from the three conditions of optimization tested, experts may point out some parameters to correct. Indeed, the contribution of experts into a modelling's design applied to a food science problematic was shown to provide a more accurate global model (Chabin et al., 2017). Hence, visual interaction with the experts could help to balance experts' knowledge with information automatically extracted from the chemical and sensory data. For instance, the comparison of the three conditions of optimization may also help to get insights on odor construction and these information might be confirmed or not by expertise.

Fuzzy logic can also cope with uncertainty assessment and management. In our approach, we did not take into account these parameters, but we previously mentioned this issue especially with regard to chemical and sensory data collected within the flavor analysis path. Moreover, expertise may also be sensitive to uncertainty so that one can go back to the experts and ask them to assess their confidence regarding their answers. From this information, a fuzzy rule for each expert could be created while integrating uncertainty parameters. This step should lead to the creation of different fuzzy rules representative of each expertise and by testing each rules, it would be possible to compare the expertise collected. If one rule predicts more accurately the intensity of a given OSA, one possibility can be then to show the results to experts to get a feedback on the rules and at the same time record insights on complex odor construction.

4. Knowledge representation vs. Fuzzy formalization

The OOPS and the integrative approach based on fuzzy logic allowed the prediction of the odor profiles of red wines. These two predictive modelling approaches were applied to predict the odor profile of two wines (PN-A and CF-A wines) selected from the initial set of 16 wines. The results of the two modelling approaches can be compared against the experimental sensory evaluation for these two wines.

The OOPS approach failed to predict the OSA Cut-grass, Leather, Toasty, Vanilla, Violet, and Woody for both wines, namely 6 OSA among 15. The fuzzy logic based integrative approach failed to predict the Cherry-related OSA (Cherry cooked, Cherry fresh, and Cherry stone) and the OSA Violet for both wines, that is 4 OSA among 15. This last approach also failed to predict the OSA Bell pepper, Blackcurrant bud, and Prune for the CF-A wine only.

In both cases, the OSA Violet was not predicted because in the OOPS formalism, no identified molecules in the wines carried the OD that are part of the OQ violet, which is directly linked to the OSA Violet since Violet is a simple OSA. One possibility to explain this issue is that the molecules related to the Violet odor were among the not identified ones. It would be possible to test this hypothesis by considering the descriptors proposed by the panel during the GC-O analysis.

Concerning the other OSA, contrasting results were obtained following the two modelling pathways. In the integrative approach based on fuzzy logic, no molecule in the CF-A wine was linked to the OQ Lactonic or to the OQ Vanilla. However, according to experts' knowledge, the OQ Lactonic is integrated in the fuzzy rules to predict the OSA Bell pepper and Prune whereas the OQ Vanilla is integrated in the OSA Blackcurrant bud. Hence with a lack of these OQ in the CF-A wine, the prediction of the three OSA was not possible. One may wonder why these OSA were nevertheless predicted in the CF-A wine with the OOPS approach. In fact, the OOPS relied on expertise but followed an additive process, which is not the case for the fuzzy experts' rules. Thus, even if one OQ among an OQ set contributing to an OSA was missing, the respective OSA can be predicted in the OOPS approach if the other OQ were present. Therefore, to increase the fuzzy logic based approach flexibility, it would be interesting to consider the OOPS additive approach only in the case of non-predicted values.

Regarding the OSA Cherry (Cherry cooked, Cherry fresh, and Cherry stone), the reasons explaining the differences in predicted values following the two approaches are less clear yet. One possible explanation would be that the profiles of the wines appeared to be closer to the sensory data with the OOPS prediction compared to the fuzzy prediction because the outputs of both models are not on the same scale: the OOPS predicted a proportion of OSA whereas the fuzzy rules allowed the prediction of the intensity of each OSA.

Finally, the comparison between the two modelling approaches was performed on only two wines of the sample set. Indeed, the OSA profile prediction from the OOPS was only made

for two wines out of the 16. One perspective is to compute the predicted profiles for all the wines in order to compare the prediction from the OOPS and the integrative fuzzy model on the whole sample set and thus to compare the results using statistical tests (PCA, HCPC, and RV coefficients).

5. Rules derived from expertise

The expertise knowledge integrated in the modelling approaches was collected from 5 flavorists. Experts from other domain can be interviewed to increase the representativity and/or genericity of the knowledge embedded in both the OOPS and the fuzzy formalization.

On the one hand, chemical scientists have knowledge about odorant's mixtures perception. Indeed, some studies about odorants' mixtures identified combination rules. For instance, Ferreira et al. (2016) studied 6 odorants' mixtures and their impact on the sensory perceptions of wines. In this case, odorants were grouped after their chemical characteristics (norisoprenoids, branched acids, enolones, branched ethyl esters, major alcohol, and oak related compounds). By considering this type of expertise, molecules could be gathered according to their chemical characteristics before the modelling step in order to balance the relative weights of their semantic descriptors. One the other hand, sensory scientists or food-domain experts rely on specific semantic descriptors regarding the food product they have to characterize. For example, lists of standardized terminology are published (caramel: Paravisini et al., 2014; wine: Noble et al., 1987). Such knowledge may also be advantageously exploited.

However, each expertise relies on specific and different corpuses (e.g. chemical names, specific sensory terms) and thus the combination of expertise from different domains should be done with caution and likely requires further development. For instance, it would be necessary to verify the knowledge aggregation in an interactive manner with the experts involved in order to validate the expertise formalization. The main outcome of such approach would be to facilitate communication among experts from different domains regarding odor mixtures perception.

CONCLUSION

Within this PhD thesis we demonstrated the possibility to predict the odor sensory profile of complex odorants' mixtures on the basis of their chemical composition. We applied our research on real food odor products being red wines. Our challenge was to integrate expert knowledge in our modelling approach and levered some bottlenecks of this domain. Results have shown the interest of such an approach, but also the complexity and multidisciplinary skills required to aggregate heterogeneous data.

This work brings innovation by its ability to predict the odor of a complex food product using semantic attributes associated to quantitative dimension. Indeed, to the best of our knowledge, the only previous work performed on odorants' mixtures focused on predicting perceptual similarity to avoid biases introduced by semantic information. To overcome this issue, we relied on mixtures' perception rules by integrating expertise into our modelling.

Altogether, this predictive work can contribute to industrial applications since the prediction of odor profiles can guide the formulation of aromas or food products. For instance, the developed models might be valuable in the assemblage of wines, i.e. a wine-making technique involving the blending of different wines prior to bottling. The key behind assemblage is to know which wines can be combined in order to obtain a targeted and balanced flavor. If the chemical composition of the wines is known, their respective odor profiles can be predicted as well as their profile if blended.

Concerning food flavor analysis, the predictive tools might be suitable to simulate omission experiments and thus might contribute to avoid fastidious experiments and thus to save time. In that sense, this work may improve the efficiency of the food flavor key volatile compound identification.

Concerning fundamental research, since the modelling approaches are "grey box", it is possible to extract rules from the expert system, which may increase our understanding of the complex integration processes that govern complex odor mixture perception.

BIBLIOGRAPHY

A

Abe, H, Kanaya, S, Komukai, T, Takahashi, Y & Sasaki, S-I. 1989. Systemization of semantic descriptions of odors. Analytica chimica acta, 239:73-85. doi:10.1016/S0003-2670(00)83837-4.

Abraham, MH, Sánchez-Moreno, R, Cometto-Muñiz, JE & Cain, WS. 2012. An algorithm for 353 odor detection thresholds in humans. Chemical senses, 37(3):207-218. doi:10.1093/chemse/bjr094.

Acree, TE, Barnard, J & Cunningham DG. 1984. A procedure for the sensory analysis of gas chromatographic effluents. Food chemistry, 14:273-286.

Allais, I, Perrot, N, Curt, C & Trystram, G. 2007. Modelling the operator know-how to control sensory quality in traditional processes. Journal of food engineering, 83(2):156-166. doi:10.1016/j.jfoodeng.2007.02.016.

Araneda, RC, Kini, AD & Firestein, S. 2000. The molecular receptive range of an odorant receptor. Nature neuroscience, 3(12):1248-1255. doi:10.1038/81774.

Arctander, S. 1969. Perfume and flavor chemicals, Volumes I and II. Montclair, NJ, USA.

Armelin-Correa, LM & Malnic, B. 2017. Combining in vivo and in vitro approaches to identify human odorant receptors responsive to food odorants. Journal od agriculture and food chemistry, 66(10):2214-2218. doi:10.1021/acs.jafc.6b04998.

Arn, H & Acree, TE. 1998. Flavornet: A database of aroma compounds based on odor potency in natural products. Developments in food science, 40:27. doi:10.1016/S0167-4501(98)80029-0.

Atanasova, B, Thomas-Danguin, T, Langlois, D, Nicklaus, S, Chabanet, C & Etiévant, P. 2005. Perception of wine fruity and woody notes: Influence of peri-threshold odorants. Food quality and preference, 16(6):504-510. doi:10.1016/j.foodqual.2004.10.004.

Audouin, V, Bonnet, F, Vickers, ZM & Reineccius, GA. 2001. Limitation in the use of odor activity values to determine important odorants in foods. In: Gas chromatography-olfactometry: The state of the art, eds Leland, JV, Schieberle, P, Buettner, A & Acree, TE. American chemical society, Washington, DC, USA. 156-171.

Aznar, M, López, R, Cacho, JF & Ferreira, V. 2001. Identification and quantification of impact odorants of aged red wines from Rioja. GC-olfactometry, quantitative GC-MS, and odor evaluation of HPLC fractions. Journal of agricultural and food chemistry, 49(6):2924-2929. doi:10.1021/jf001372u.

Aznar, M, López, R, Cacho, J & Ferreira, V. 2003. Prediction of aged red wine aroma properties from aroma chemical composition. Lartial least squares regression models. Journal of agricultural and food chemistry, 51(9):2700-2707. doi:10.1021/jf026115z.

B

Ballester, J, Patris, B, Symoneaux, R & Valentin, D. 2008. Conceptual vs? perceptual wine spaces: Does expertise matter? Food quality and preference, 19(3):267-276. doi:10.1016/j.foodqual.2007.08.001.

Banerjee, R, Tudu, B, Bandyopadhyay, R & Bhattacharyya, N. 2016. A review on combined odor and taste sensor systems. Journal of food engineering, 190:10-21.

Barba, C, Thomas-Danguin, T & Guichard, E. 2017. Comparison of stir bar sorptive extraction in the liquid and vapour phases, solvent-assisted flavour evaporation and headspace solid-phase microextraction for the (non)-targeted analysis of volatiles in fruit juice. LWT – Food science and technology, 85(B):334-344. doi:10.1016/j.lwt.2016.09.015.

Barkat-Defradas, M & Motte-Florac, E. 2016. Verbalization of olfactory perception. In: Words for odours: Language skills and cultural insights, eds Barkat-Defradas, M & Motte-Florac, E. Cambridge scholars publishing, Newcastle upon Tyne, UK. 1-22.

Bell, GA, Laing, DG & Panhuber, H. 1987. Odour mixture suppression: Evidence for a peripheral mechanism in human and rat. Brain research, 426(1):8-18. doi:10.1016/0006-8993(87)90419-7.

Berglund, B, Berglund, U & Lindvall, T. 1976. Psychological processing of odor mixtures. Psychological Review, 83(6):432-441. doi:10.1037/0033-295X.83.6.432.

Berglund, B & Olsson, MJ. 1993a. Odor-intensity interaction in binary mixtures. Journal or experimental psychology: Human perception and performance, 19(2):302-314. doi:10.1037/0096-1523.19.2.302.

Berlung, B & Olsson, MJ. 1993b. Odor-intensity interaction in binary and ternary mixtures. Perception & psychophysics, 53(5), 475–482. doi:10.3758/BF03205195.

Boccorh, RK & Paterson, A. 2002. An artificial neural network model for predicting flavour intensity in blackcurrant concentrates. Food quality and preference, 13(2):117-128. doi:10.1016/S0950-3293(01)00072-6.

Boesveldt, S & de Graaf, S. 2017. The differential role of smell and taste for eating behavior. Perception, 46(3-4):307-319. doi:10.1177/0301006616685576.

Boyle, JA, Djordjevic, Olssons, MJ, Lundström, JN & Jones-Gotman, M. 2009. The human brain distinguishes between single odorants and binary mixtures. Cerebral cortex, 19(1):66-71. doi:10.1093/cercor/bhn058.

Bratolli, M, Cisternino, E, Dambruoso, PR, de Gennaro, G, Giungato, P, Mazzone, A, Palmisani, J & Tutino, M. 2013. gas chromatography analysis with olfactometric detection (GC-O) as a useful methodology for chemical characterization of odorous compounds. Sensors, 13:16759-16800. doi:10.3390/s131216759.

Breslin, PAS & Huang, L. 2006. Human taste: Peripheral anatomy, taste transduction, and coding. In: Taste and smell: An update, eds Hummel, T & Welge-Lüssen, A. Karger AG, Basel, Switzerland. 152-190.

Briand, L & Salles, C. 2016. Taste perception and integration. In: Flavor: From food to behaviors, wellbeing and health, eds Etiévant, E, Guichard, E, Salles, C & Voilley, A. Woodhead publishing. 101-119.

Brockhoff, PB, Schlich, P & Skovgaard, I. 2015. Taking individual scaling differences into account by analyzing profile data with the Mixed Assessor Model. Food quality and preference, 39:156-166. doi:10.1016/j.foodqual.2014.07.005.

Brodin, M, Laska, M & Olsson, MJ. 2009. Odor interaction between bourgeonal and its antagonist undecanal. Chemical senses, 34(7):625-630. doi:10.1093/chemse/bjp044.

Buck, L & Axel, R. 1991. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. Cell, 65(1):175-187. doi:10.1016/0092-8674(91)90418-X.

Buettner, A, Beer, A, Hannig, C & Settles, M. 2001. Observation of the swallowing process by application of videofluoroscopy and real-time magnetic resonance imaging - consequences for retronasal aroma stimulation. Chemical senses, 26(9):1211-1219. doi:10.1093/chemse/26.9.1211.

Burdock, GA. 2010. Fenaroli's handbook of flavor ingredients: Sixth edition. CRC press, Boca Raton, FL, USA.

Bushdid, C, Magnasco, MO, Vosshall, LB & Keller, A. 2014. Humans can discriminate more than 1 trillion olfactory stimuli. Science, 343(6177):1370-1372. doi:10.1126/science.1249168.

C

Cain, WS. 1975. Odor intensity: Mixtures and masking. Chemical senses, 1:339-352. doi:10.1093/chemse/1.3.339.

Cain, WS. 1979. To know with the nose: Keys to odor identification. Science, 203(4379):467-470. doi:10.1126/science.760202.

Cain, WS & Drexler, M. 1974. Scope and evaluation of odor counteraction and masking. Annals of the New York academy of sciences, 237:427-439. doi:10.1111/j.1749-6632.1974.tb49876.x.

Cameleyre, M, Lytra, G, Tempere, S & Barbe, J-C. 2015. Olfactory impact of higher alcohols on red wine fruity ester aroma expression in model solution. Journal of agricultural and food chemistry, 63(44):9777-9788. doi:10.1021/acs.jafc.5b03489.

Campo, E, Ferreira, V, Escudero, A & Cacho, J. 2005. Prediction of the wine sensory properties related to grape variety from dynamic-headspace gas chromatography-olfactometry data. Journal of agricultural and food chemistry, 53(14):5682-5690. doi:10.1021/jf047870a.

Cancilla, JC, Wang, SC, Díaz-Rodríguez, P, Matute, G, Cancilla, JD, Flynn, D & Torrecilla, JS. 2014. Linking chemical parameters to sensory panel results through neural networks to distinguish olive oil quality. Journal of agricultural and food chemistry, 62(44):10661-10665. doi:10.1021/jf503482h.

Castro, JB, Ramanathan, A & Chennubhotla, CS. 2013. Categorical Dimensions of Human odor descriptor space revealed by non-negative matrix factorization. Plos one, 8(9):e73289. doi:10.1371/journal.pone.0073289.

Cellier, J-M, Eyrolle, H & Marine, C. 1997. Expertise in dynamic environments. Ergonomics, 40(1):28-50. doi:10.1080/001401397188350.

Chabin, T, Barnabe, M, Boukhelifa, N, Fonseca, F, Tonda, A, Velly, H, Lemaitre, B, Perrot, N & Lutton, E. 2017. LIDeOGraM : An interactive evolutionary modelling tool. 13th Biennial International Conference on Artificial Evolution (EA-2017), Paris, France. 25-27 October.

Chambers, E & Kopple, K. 2013. Associations of volatile compounds with sensory aroma and flavor: the complex nature of flavor. Molecules, 18:4887-4905. doi:10.3390/molecules18054887.

Chaput, MA, El Mountassir, F, Atanasova, B, Thomas-Danguin, T, Le Bon, AM, Perrut, A, Ferry, B & Duchamp-Viret, P. 2012. Interactions of odorants with olfactory receptors and receptor neurons match the perceptual dynamics observed for woody and fruity odorant mixtures. European journal of neuroscience, 35(4):584-597. doi:10.1111/j.1460-9568.2011.07976.x.

Chastrette, M. 1998. Data management in olfaction studies. SAR and QSAR in environmental research, 8(3-4):157-181. doi:10.1080/10629369808039139.

Chastrette, M. 2002. Classification of odors and structure-odor relationships. In: Olfaction, taste and cognition eds Rouby, C, Schaal, B, Dubois, D, Gervais, R & Holley, A. Cambridge university press, New York, NY, USA. 100-116.

Chastrette, M, de Saint Laumer, JY. 1991. Structure-odor relationships using neural networks. European journal of medicinal chemistry, 26(8):829-833. doi:10.1016/0223-5234(91)90010-K.

Chastrette, M, de Saint Laumer, J-Y & Sauvegrain, P. 1991. Analysis of a system of description of odors by means of four different multivariate statistical methods. Chemical senses, 16(1):81-93. doi:10.1093/chemse/16.1.81.

Chastrette, M, Thomas-Danguin, T & Rallet, E. 1998. Modelling the human olfactory stimulus-response function. Chemical senses, 23(2):181-196. doi:10.1093/chemse/23.2.181.

Chaudhari, N & Roper, SD. 2010. The cell biology of taste. Journal of cell biology, 190(3):285-296. doi:10.1083/jcb.201003144.

Chee-Ruiter, CWJ & Bower, JM. 1998. Representing odor quality space: A perceptual framework for olfactory processing. In: Computational Neuroscience: Trends in research, ed Bower, JM. Plenum press, New York, NY, USA. 591-598. doi:10.1007/978-1-4615-4831-7_98.

Chi, MTH, Feltovich, PJ & Glaser, R. 1981. Categorization and representation of physics problems by experts and novices. Cognitive science, 5(2):121-152. doi:10.1207/s15516709cog0502_2.

Chrea, C, Valentin, D, Sulmont-Rossé, C, Ly Mai, H, Hoang Nguyen, D & Abdi, H. 2004. Culture and odor categorization: Agreement between cultures depends upon the odors. Food quality and preference, 15:669-679. doi:10.1016/j.foodqual.2003.10.005.

Coureaud, G, Gibaud, D, Le Berre, E, Schaal, B & Thomas-Danguin, T. 2011. Proportion of odorants impacts the configural versus elemental perception of a binary blending mixture in newborn rabbits. Chemical Senses, 36(8):693-700. doi:10.1093/chemse/bjr049.

Croijmans, I & Majid, A. 2016. Not all flavor expertise is equal: The language of wine and coffee experts. Plos one, 11(6):e0155845. doi:10.1371/journal.pone.0155845.

Czerny, M, Christlbauer, M, Fischer, A, Granvogl, M, Hammer, M, Hartl, C, Hernandez, NM & Schieberle, P. 2008. Re-investigation on odour thresholds of key food aroma compounds and development of an aroma language based on odour qualities of defined aqueous odorant solutions. European food research and technology, 228(2):265-273. doi:10.1007/s00217-008-0931-x.

D

d'Acampora Zellner, B, Dugo, P, Dugo, G & Mondello, L. 2008. Gas chromatography-olfactometry in food flavour analysis. Journal of chromatography A, 1186:123-143. doi:10.1016/j.chroma.2007.09.006.

Da Costa, NC & Eri, S. 2009. Identification of aroma chemicals. In: Chemistry and Technology of Flavors and Fragrances, ed Rowe, DJ. CRC press, Boca Raton, FL, USA.Boca Raton, Florida: LLC. 12-34.

de-la-Fuente-Blanco, A, Sáenz-Navajas, M-P & Ferreira, V. 2016. On the effects of higher alcohols on red wine aroma. Food chemistry, 210:107-114. doi:10.1016/j.foodchem.2016.04.021.

de-la-Fuente-Blanco, A, Sáenz-Navajas, M-P & Ferreira, V. 2017. Levels of higher alcohols inducing aroma changes and modulating experts' preferences in wine model solutions. Australian journal of grape and wine research, 23(2):162-169. doi:10.1111/ajgw.12273.

Delahunty, CM, Eyres, G & Dufour, J-P. 2006. Gas chromatography-olfactometry. Journal of separation science, 29(14):2107-2125. doi:10.1002/jssc.200500509.

de la Presa Owens, C, Schlich, P, Wada, K. & Noble, AC. 1998. Using sensory and instrumental data to interpret the effect of storage at elevated temperatures on aroma of Chardonnay wines. Annals of the New York academy of sciences, 855:854-859. doi:10.1111/j.1749-6632.1998.tb10675.x.

Déléris, I, Guichard, E & Salles, C. 2012. Libération des composes d'arôme et des composes sapides. In: Texture et flaveur des aliments, eds Guichard, E, Genot, C & Voilley, A. Educagri éditions, Dijon, France. 203-231.

Delime, P, O'Mahony, K, Desforges, N, Taylor, AJ & Hort, J. 2016. Comparing the relative sensitivity of ortho- and retronasal perception of a strawberry flavour model using omission testing. Flavour and fragrance journal, 31:377-384. doi:10.1002/ffj.3326.

Dernoncourt, F. 2011. Introduction à la logique floue. http://franck-dernoncourt.developpez.com/tutoriels/algo/introduction-logique-floue/

Doty, RL, Brugger, WE, Jurs, PC, Orndorff, MA, Snyder PJ & Lowry, LD. 1978. Intranasal trigeminal stimulation from odorous volatiles: Psychometric responses from anosmic and normal humans. Physiology and behavior, 20(2):175-185. doi:10.1016/0031-9384(78)90070-7.

Dravnieks, A. 1985. Atlas of odor character profiles. ASTM, Philadelphia, PA, USA.

Dubois, D. 2006. Des catégories d'odorants à la sémantique des odeurs : Une approche cognitive de l'olfaction. Terrain, 47:89-106. doi:10.4000/terrain.4263.

Dubois, D & Prade, H. 1980. Fuzzy sets and systems: Theory and applications. Academic press, New York, NY, USA.

Duchamp, A, Revial, MF, Holley, A & Mac Leod, P. 1974. Odor discrimination by frog olfactory receptors. Chemical senses, 1(2):213-233. doi:10.1093/chemse/1.2.213.

Duchamp-Viret, P, Duchamp, A & Chaput, MA. 2003. Single olfactory sensory neurons simultaneously integrate the components of an odour mixture. European journal of neuroscience, 18:2690-2696. doi:10.1111/j.1460-9568.2003.03001.x.

E - F

Edwards, PA & Jurs, PC. 1989. Correlation of odor intensities with structural properties of odorants. Chemical senses, 14(2):281-291. doi:10.1093/chemse/14.2.281.

Escudero, A, Gogorza, B, Melús, MA, Ortín, N, Cacho, J & Ferreira, V. 2004. characterization of the aroma of a wine from maccabeo. Key role played by compounds with low odor activity values. Journal of agricultural and food chemistry, 52(11):3516-3524. doi:10.1021/jf035341l.

Ferdenzi, C, Joussain, P, Digard, B, Lunneau, L, Djordjevic, J & Bensafi, M. 2016. individual differences in verbal and non-verbal affective responses to smells: Influence of odor label across cultures. Chemical senses, 42(1):37-46. doi:10.1093/chemse/bjw098.

Ferreira, V, Ortín, N, Escudero, A, López, R & Cacho, J. 2002. Chemical characterization of the aroma of Grenache rosé wines: Aroma extract dilution analysis, quantitative determination, and sensory reconstitution studies. Journal of agricultural and food chemistry, 50(14):4048-4054. doi:10.1021/jf0115645.

Ferreira, V. 2012a. Revisiting psychophysical work on the quantitative and qualitative odour properties of simple odour mixtures: A flavour chemistry view. Part 1: Intensity and detectability. A review. Flavour and fragrance journal, 27:124-140. doi:10.1002/ffj.2090.

Ferreira, V. 2012b. Revisiting psychophysical work on the quantitative and qualitative odour properties of simple odour mixtures: A flavour chemistry view. Part 2: Qualitative aspects. A review. Flavour and fragrance journal, 27:201-215. doi:10.1002/ffj.2091.

Ferreira, V, Sáenz-Navajas, MP, Campo, E, Herrero, P, de la Fuente, A & Fernández-Zurbano, P. 2016. Sensory interactions between six common aroma vectors explain four main red wine aroma nuances. Food chemistry, 199:447-456. doi:10.1016/j.foodchem.2015.12.048.

Fisher, C & Scott, TR.1997. Food flavours: Biology and Chemistry. RSC, Cambridge, UK. 1-14.

Fondberg, R, Lundström, JN, Blöch, M, Olsson, MJ & Seubert, J. 2018. Multisensory flavor perception: The relationship between congruency, pleasantness, and odor referral to the mouth. Appetite, 125:244-252. doi:10.1016/j.appet.2018.02.012.

Friedman, L & Miller JG. 1971. Odor incongruity and chirality. Science, 172:1044-1046. doi:10.1126/science.172.3987.1044.

G

Goldberg, EM, Wang, K, Goldberg, J & Aliani, M. 2018. Factors affecting the ortho- and retronasal perception of flavors: A review. Critical review in food science and nutrition, 58(6):913-923. doi:10.1080/10408398.2016.1231167.

González Alvarez, M, González –Barreiro, C, Cancho-Grande, B & Simal-Gándara, J. 2011. Relationships between Godello white wine sensory properties and its aromatic fingerprinting obtained by GC-MS. Food chemistry, 129(3):890-898. doi:10.1016/j.foodchem.2011.05.040.

Goyert, HF, Frank, ME, Gent, JF & Hettinger, TP. 2007. Characteristic component odors emerge from mixtures after selective adaptation. Brain research bulletin, 72(1):1-9. doi:10.1016/j.brainresbull.2006.12.010.

Grosch, W. 2001. Evaluation of the key odorants of foods by dilution experiments, aroma models and omission. Chemical senses, 26(5):533-545. doi:10.1093/chemse/26.5.533.

Gross-Isserof, R & Lancet, D. 1988. Concentration-dependent changes of perceived odor quality. Chemical senses, 13(2):191-204. doi:10.1093/chemse/13.2.191.

Grossman, KJ, Mallik, AK, Ross, J, Kay, LM & Issa, NP. 2008. Glomerular activation patterns and the perception of odor mixtures. European journal of neuroscience, 27(10):2676-2685. doi:10.1111/j.1460-9568.2008.06213.x.

Gruber, TR. 1993. A translation approach to portable ontology specifications. Knowledge acquisition, 5(2):199-220. doi:10.1006/knac.1993.1008.

H

Haddad, R, Medhanie, A, Roth, Y, Harel, D & Sobel, N. 2010. Predicting odor pleasantness with an electronic nose. Plos computational biology, 6(4):e1000740. doi:10.1371/journal.pcbi.1000740.

Hayes, JE & Pickering, GJ. 2012. Wine expertise predicts taste phenotype. American journal of enology and viticulture, 63(1):80-84. doi:10.5344/ajev.2011.11050.

Heilmann, S & Hummel, T. 2004. A new method for comparing orthonasal and retronasal olfaction. Behavioral neuroscience, 118(2):412-419. doi:10.1037/0735-7044.118.2.412.

Hornung, DE & Enns, MP. 1986. The contributions of smell and taste to overall intensity: A model. Perception and psychophysics, 39(6):385-391.doi:10.3758/BF03207066.

Hudelot, C. 2014. Quelques contributions à l'interprétation et à l'annotation sémantiques d'images : Apport des approches logiques et des ontologies et de la prise en compte de l'imperfection. Mémoire présenté pour obtenir le titre d'habilitation à diriger des recherches de l'université Paris-Sud XI.

Hudson, R. 2000. Odor and odorant: A terminological clarification. Chemical senses, 25(6):693. doi:10.1093/chemse/25.6.693.

Hughson, AL & Boakes, RA. 2002. The knowing nose: the role of knowledge in wine expertise. Food quality and preference, 13:463-472. doi:10.1016/S0950-3293(02)00051-4.

I

Iatropoulos, G, Herman, P, Lansner, A, Karlgren, J, Larsson, M & Olofsson, JK. 2018. The language of smell: Connecting linguistic and psychophysical properties of odor descriptors. Cognition, 178:37-49. doi:10.1016/j.cognition.2018.05.007.

Ibanescu, L, Dibie, J, Dervaux, S, Guichard, E & Raad, J. 2016. PO² - A process and observation ontology in food science. Application to Dairy Gels. Metadata and Semantics Research - 10th International Conference, Gottingen, Germany, 22-25 November. doi:10.1007/978-3-319-49157-8_13.

Ickes, C & Cadwallader, K. 2017. Effects of ethanol on flavor perception in alcoholic beverages. Chemosensory perception, 10(4):119-134. doi:10.1007/s12078-017-9238-2.

Ioannou, I, Perrot, N, Hossenlopp, J, Mauris, G & Trystram, G. 2002. The fuzzy set theory: A helpful tool for the estimation of sensory properties of crusting sausage appearance by a single expert. Food quality and preference, 13:589-595. doi:10.1016/S0950-3293(02)00045-9.

Ishii, A, Roudnitzky, N, Béno, N, Bensafi, M, Hummel, T, Rouby, C & Thomas-Danguin, T. 2008. Synergy and masking in odor mixtures: An electrophysiological study of orthonasal vs. retronasal perception. Chemical senses, 33(6):553-561. doi:10.1093/chemse/bjn022.

J

Jacquier, V, Pick, H & Vogel, H. 2006. Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17-40 comprising structurally related compounds. Journal of neurochemistry, 97:537-544. doi:10.1111/j.1471-4159.2006.03771.x.

James, N, Todorov, K & Hudelot, C. 2010. Ontology matching for the semantic annotation of images. International Conference on Fuzzy Systems, Barcelona, Spain, 18-23 July. doi:10.1109/FUZZY.2010.5584354.

Jaubert, J-N, Tapiero, C & Dore, J-C. 1995. The field of odors: Toward a universal language for odor relationships. Perfumer and flavorist, 20:1-16.

Jiang, Y, Gong, N & Matsunami, H. 2014. Astringency: A more stringent definition. Chemical senses, 39(6):467-469. doi:10.1093/chemse/bju021.

Jinks, A & Laing, DG. 2001. The analysis of odor mixtures by humans: Evidence for a configurational process. Physiology and behavior, 72(1-2):51-63. doi:10.1016/S0031-9384(00)00407-8.

K

Kay, LM, Lowry, CA & Jacobs, HA. 2003. Receptor contributions to configural and elemental odor mixture perception. Behavioral neuroscience, 117(5):1108-1114. doi:10.1037/0735-7044.117.5.1108.

Kay, LM, Crk, T & Thorngate, J. 2005. A redefinition of odor mixture quality. Behavioral neuroscience, 119(3):726-733. doi:10.1037/0735-7044.119.3.726.

Keller, A, Hempstead, M, Gomez, IA, Gilbert, AN & Vosshall, LB. 2012. An olfactory demography of a diverse metropolitan population. BMC Neuroscience, 13:122. doi:10.1186/1471-2202-13-122.

Keller, A & Vosshall, LB. 2016. Olfactory perception of chemically diverse molecules. BMC Neuroscience, 17:55. doi:10.1186/s12868-016-0287-2.

Keller, A, Gerkin, RC, Guan, Y, Dhurandhar, A, Turu, G, Szalai, B, Mainland, JD, Ihara, Y, Yu, CV, Wolfinger, R, Vens, C, Schietgat, L, De Grave, K, Norel, R, DREAM Olfaction Prediction Consortium, Stolovitzky, G, Cecchi, GA, Vosshall, LB & Meyer, P. 2017. Predicting human olfactory perception from chemical features of odor molecules. Science, 355(6327):820-826. doi:10.1126/science.aal2014.

Kermen, F, Chakirian, A, Sezille, C, Joussain, P, Le Goff, G, Ziessel, A, Chastrette, M, Mandairon, N, Didier, A, Rouby, C & Bensafi, M. 2011. Molecular complexity determines the number of olfactory notes and the pleasantness of smells. Scientific reports, 1:206. doi:10.1038/srep00206.

Khan, RM, Luk, C-H, Flinker, A, Aggarwal, A, Lapid, H, Haddad, R & Sobel, N. 2007. Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. Journal of neuroscience, 27(37):10015-10023. doi:10.1523/JNEUROSCI.1158-07.2007.

Kissi, M, Ramdani, M, Tollabi, M & Zakarya, D. 2004. Determination of fuzzy logic membership functions using genetic algorithms: Application to structure-odor modeling. Journal of molecular modeling, 10:335-341. doi:10.1007/s00894-004-0200-2.

Kumar, R, Kaur, R, Auffarth, B & Bhondekar, AP. 2015. Understanding the odour spaces: A step towards solving olfactory stimulus-percept problem. Plos one, 10(10): e0141263. doi:10.1371/journal.pone.0141263.

Kuo, Y-L, Pangborn, RM & Noble, AC. 1993. Temporal patterns of nasal, oral, and retronasal perception of citral and vanillin and interaction of these odourants with selected tastants. International journal of food science and technology, 28:127-137. doi:10.1111/j.1365-2621.1993.tb01258.x.

L

Laffort, P & Dravnieks, A. 1982. Several models of suprathreshold quantitative olfactory interaction in humans applied to binary, ternary and quaternary mixtures. Chemical senses, 7(2):153-174. doi:10.1093/chemse/7.2.153.

Laing, DG. 1994. Perceptual odour interactions and objective mixture analysis. Food quality and preference, 5:75-80.

Laing, DG & Francis, GW. 1989. The capacity of humans to identify odors in mixtures. Physiology and behavior, 46(5):809-814. doi:10.1016/0031-9384(89)90041-3.

Laing, D & Jinks, A. 1996. Flavour perception mechanisms. Trends in food science and technology, 7(12):387-389. doi:10.1016/S0924-2244(96)10049-2.

Lapid, H, Harel, D & Sobel, N. 2008. Prediction models for the pleasantness of binary mixtures in olfaction. Chemical senses, 33(7):599-609. doi:10.1093/chemse/bjn026.

Lapis, TJ, Penner, MH & Lim, J. 2016. Humans can taste glucose oligomers independent of the hT1R2/hT1R3 sweet taste receptor. Chemical senses, 41(9):755-762. doi:10.1093/chemse/bjw088.

Lawless, H & Engen, T. 1977. Associations to odors: Interference, mnemonics, and verbal labeling. Journal of experimental psychology: Human learning and memory, 3(1):52-59. doi:10.1037/0278-7393.3.1.52.

Lawless, HT & Heymann, H. 2010. Sensory evaluation of food: Principles and practices. Second edition. Springer Science+Business Media, New York, USA. 227-257.

Le Berre, E, Atanasova, B, Langlois, D, Etiévant, P & Thomas-Danguin, T. 2007. Impact of ethanol on the perception of wine odorant mixtures. Food quality and preference, 18(6):901-908. doi:10.1016/j.foodqual.2007.02.004.

Lee, S-J, & Noble, AC. 2003. Characterization of odor-active compounds in Californian Chardonnay wines using GC-olfactometry and GC-mass spectrometry. Journal of agricultural and food chemistry, 51(27):8036-8044. doi:10.1021/jf034747v.

Le Fur, Y, Mercurio, V, Moio, L, Blanquet, J & Meunier JM. 2003. A new approach to examine the relationships between sensory and gas chromatography-olfactometry data using generalized procrustes analysis applied to six French Chardonnay wines. Journal of agricultural and food chemistry, 51:443-452. doi:10.1021/jf0205458.

Li, H, Panwar, B, Omenn, GS & Guan, Y. 2018. Accurate prediction of personalized olfactory perception from large-scale chemoinformatic features. Gigascience, 7(2):1-11. doi:10.1093/gigascience/gix127.

Lin, DY, Shea, SD & Katz, LC. 2006. Representation of natural stimuli in the rodent main olfactory bulb. Neuron, 50(6):937-949. doi:10.1016/j.neuron.2006.03.021.

Lindqvist, A, Höglund, A & Berglund, B. 2012. The role of odour quality in the perception of binary and higher-order mixtures. Perception, 41(11):1373-1391. doi:10.1068/p7267.

Linko, S. 1998. Expert systems – What can they do for the food industry? Trends in food science and technology, 9(1):3-12. doi:10.1016/S0924-2244(97)00002-2.

Livermore, A & Laing, DG. 1998. The influence of odor type on the discrimination and identification of odorants in multicomponent odor mixtures. Physiology and behavior, 65(2):311-320. doi:10.1016/S0031-9384(98)00168-1.

Lorrain, B, Ballester, J, Thomas-Danguin, T, Blanquet, J, Meunier, JM & Le Fur, Y. 2006. Selection of potential impact odorants and sensory validation of their importance in typical Chardonnay wines. Journal of agricultural and food chemistry, 54(11):3973-3981. doi:10.1021/jf052350g.

Loufti, A, Coradeschi, S, Mani, GK, Shankar, P & Rayappan, JPP . 2015. Electronic noses for food quality: A review. Journal of food engineering, 144:103-111. doi:10.1016/j.jfoodeng.2014.07.019.

Luebke, W. 1980. The good scents company. Available at: http://www.thegoodscentscompany.com/index.html

Lytra, G, Tempere, S, Le Floch, A, de Revel, G & Barbe J-C. 2013. Study of sensory interactions among red wine fruity esters in a model solution. Journal of agricultural and food chemistry, 61(36):8504-8513. doi:10.1021/jf4018405.

M

Macfie, JHJ & Hedderley, D. 1993. Current practice in relating sensory perception to instrumental measurements. Food quality and preference, 4:41-49. doi:10.1016/0950-3293(93)90312-T.

Madalli, DP, Chatterjee, U & Dutta, B. 2017. An analytical approach to building a core ontology for food. Journal of documentation, 73(1):123-144. doi:10.1108/JD-02-2016-0015.

Mainland, JD, Li, YR, Zhou, T, Liu, WLL & Matsunami, H. 2015. Human olfactory receptor responses to odorants. Scientific data, 2:150002. doi:10.1038/sdata.2015.2.

Malnic, B, Hirono, J, Sato, T & Buck, LB. 1999. Combinatorial receptor codes for odors. Cell, 96(5):713-723. doi:10.1016/S0092-8674(00)80581-4.

Malnic, B, Godfrey, PA & Buck, LB. 2004. The human olfactory receptor gene family. Proceedings of the national academy of sciences, 101(8):2584-2589. doi:10.1073/pnas.0307882100.

Martínez-Mayorga, K, Peppard, TL, Yongye, AN, Santos, R, Giulianotti, M & Medina-Franco, JL. 2011. Characterization of a comprehensive flavor database. Journal of chemometrics, 25:550-560. doi:10.1002/cem.1399.

Maruyama, Y, Yasuda, R, Kuroda, M & Eto, Y. 2012. Kokumi substances, enhancers of basic tastes, induce responses in calcium-sensing receptor expressing taste cells. Plos one, 7(4):e34489. doi:10.1371/journal.pone.0034489.

Mattes, RD. 2010. Fat taste in humans: Is it a primary? In: Fat detection: Taste, texture, and post ingestive effects, eds Montmayeur, JP & le Coutre, J. CRC press, Boca Raton, FL, USA. 167-193.

Medjkoune, M, Harispe, S, Montmain, J, Cariou, S, Fanlo, J-L & Fiorini, N. 2016. Towards a non-oriented approach for the evaluation of odor quality. 16th International Conference, Information processing and management of uncertainty in knowledge-based systems IPMU 2016, Eindhoven, The Netherlands, 20-24 June. doi:10.1007/978-3-319-40596-4_21.

Michishita, T, Akiyama, M, Hirano, Y, Ikeda, M, Sagara, Y & Araki, T. 2010. Gas chromatography/olfactometry and electronic nose analyses of retronasal aroma of espresso and correlation with sensory evaluation by an artificial neural network. Journal of food science, 75(9):S477-S489. doi:10.1111/j.1750-3841.2010.01828.x.

Mombaerts, P, Wang, F, Dulac, C, Chao, SK, Nemes, A, Mendelsohn, M, Edmonson, J & Axel, R. 1996. Visualizing an olfactory sensory map. Cell, 87(4):675-686. doi:10.1016/S0092-8674(00)81387-2.

Monahan, K & Lomvardas, S. 2015. Monoallelic expression of olfactory receptors. Annual review of cell and developmental biology, 31:721-740. doi:10.1146/annurev-cellbio-100814-125308.

Murphy, C & Cain, WS. 1980. Taste and olfaction: Independence vs interaction. Physiology and behavior, 24(3):601-605. doi:10.1016/0031-9384(80)90257-7.

N - O

Nagai, MH, Armellin-Correa, LM & Malnic, B. 2016. Monogenic and monoallelic expression of odorant receptors. Molecular pharmacology, 90(5):633-639. doi:10.1124/mol.116.104745.

Nagata, Y. 2003. Measurement of odor threshold by triangle odor bag method. In: Odor measurement review, ed Japan ministry of the environment, Government of Japan, Japan. 118-127.

Noble, AC, Arnold, RA, Buechsenstein, J, Leach, EJ, Schmidt, JO & Stern, PM. 1987. Modification of a standardized system of wine aroma terminology. American journal of enology and viticulture, 38:143-146.

Noy, NF & McGuinness, DL. 2001. Ontology development 101: A guide to creating your first ontology. Technical report SMI-2001-0880, Stanford medical informatics.

Nozaki, Y & Nakamoto, T. 2018. Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. Plos one, 13(6):e0198475. doi:10.1371/journal.pone.0198475.

Ojha, PK & Roy, K. 2018. Chemometric modeling of odor threshold property of diverse aroma components of wine. RSC Advances, 8:4750-4760. doi:10.1039/C7RA12295K.

Oka, Y, Omura, M, Kataoka, H & Touhara, K. 2004. Olfactory receptor antagonism between odorants. The EMBO journal, 23:120-126. doi:10.1038/sj.emboj.7600032.

P

Paravisini, L, Septier, C, Moretton, C, Nigay, H, Arvisenet, G, Guichard, E & Dacremont, C. 2014. Caramel odor: Contribution of volatile compounds according to their odor qualities to caramel typicality. Food research international, 57:79-88. doi:10.1016/j.foodres.2014.01.009.

Patte, F & Laffort, P. 1979. Alternative model of olfactory quantitative interaction in binary-mixtures. Chemical senses and flavor, 4(4):267-274. doi:10.1093/chemse/4.4.267.

Patton, S & Josephson, DV. 1957. A method for determining significance of volatile flavor compounds in foods. Journal of food science, 22:316-318. doi:10.1111/j.1365-2621.1957.tb17017.x.

Perrot, N, Trystam, G, Le Guennec, D & Guely, F. 1996. Sensor fusion for real time quality evaluation of biscuit during baking. Comparison between bayesian and fuzzy approaches. Journal of food engineering, 29:301-315. doi:10.1016/0260-8774(95)00078-X.

Perrot, N, Agioux, L, Ioannou, I, Mauris, G, Corrieu, G & Trystram, G. 2004. Decision support system design using the operator skill to control cheese ripening - Application of the fuzzy symbolic approach. Journal of food engineering, 64(3):321-333. doi:10.1016/j.jfoodeng.2003.10.019.

Perrot, N, Ioannou, I, Allais, I, Curt, C, Hossenlopp, J & Trystram, G. 2006. Fuzzy concepts applied to food product quality control: A review. Fuzzy sets and systems, 157(9):1145-1154. doi:10.1016/j.fss.2005.12.013.

Perrot, N, Trelea, IC, Baudrit, C, Trystram, G & Bourgine, P. 2011. Modelling and analysis of complex food systems: State of the art and new trends. Trends in food science and technology, 22(6):304-314. doi:10.1016/j.tifs.2011.03.008.

Perrot, N & Baudrit, C. 2013. Intelligent quality control systems in food processing based on fuzzy logic. In: Robotics and automation in the food industry - Current and future technologies, ed Caldwell, DG. Woodhead Publishing. 200-225. doi: 10.1533/9780857095763.1.200.

Peterlin, Z, Firestein, S & Rogers, ME. 2014. The state of the art of odorant receptor deorphanization: A report from the orphanage. The journal of general physiology, 143(5):527-542. doi:10.1085/jgp.201311151.

Plutowska, B & Wardencki, W. 2008. Application of gas chromatography-olfactometry (GC-O) in analysis and quality assessment of alcoholic beverages - A review. Food chemistry, 107:449-463. doi:10.1016/j.foodchem.2007.08.058.

Pollien, P, Ott, A, Montigon, F, Baumgartner, M, Muñoz-Box, R & Chaintreau, A. 1997. Hyphenated headspace-gas chromatography-sniffing technique: Screening of impact odorants and quantitative aromagram comparisons. Journal of agricultural and food chemistry, 45:2630-2637. doi:10.1021/jf960885r.

Q - R

Ressler, KJ, Sullivan, SL & Buck LB. 1994. A molecular dissection of spatial patterning in the olfactory system. Current opinion in neurobiology, 4(4):588-596. doi:10.1016/0959-4388(94)90061-2.

Rinaldi, A. 2007. The scent of life. EMBO reports, 8:629-633. doi:10.1038/sj.embor.7401029.

Rodríguez-Méndez, ML, De Saja, JA, González-Antón, R, García-Hernández, C, Medina-Plaza, C, García-Cabezón, C & Martín-Pedrosa, F. 2016. Electronic noses and tongues in wine industry. Frontiers in bioengineering and biotechnology, 4(81). doi:10.3389/fbioe.2016.00081.

Romagny, S, Coureaud, G & Thomas-Danguin, T. 2018. Key odorants or key associations? Insights into elemental and configural odour processing. Flavour and fragrance journal, 33:97-105. doi:10.1002/ffj.3429.

Rospars, J-P, Lansky, P, Chaput, M & Duchamp-Viret, P. 2008. Competitive and noncompetitive odorant interactions in the early neural coding of odorant mixtures. Journal of neuroscience, 28(10):2659-2666. doi:10.1523/JNEUROSCI.4670-07.2008.

Rossiter, KJ. 1996. Structure-odor relationships. Chemical reviews, 96(8):3201-3240. doi:10.1021/cr950068a.

Rouby, C, Thomas-Danguin, T, Sicard, G, Vigouroux, M, Jiang, T, Poitevineau, J & Issanchou, S. 2005. Influence of the semantic context on odor identification performance. Psychologie française, 50:225-239. doi:10.1016/j.psfr.2004.11.003.

Rozin, P. 1982. "Taste-smell confusions" and the duality of the olfactory sense. Perception and psychophysics, 31(4):397-401.

S

Salles, C, Briand, L, Brachais, L & Voilley, A. 2012. Molécules aromatisantes et sapides. In: Texture et flaveur des aliments, eds Guichard, E, Genot, C & Voilley, A. Educagri éditions, Dijon, France. 31-58.

Sanz, G, Schlegel, C, Pernollet, J-C & Briand, L. 2005. Comparison of odorant specificity of two human olfactory receptors from different phylogenetic classes and evidence for antagonism. Chemical senses, 30(1):69-80. doi:10.1093/chemse/bji002.

Scott, SM, James, D & Ali, Z. 2006. Data analysis for electronic nose systems. Microchimica acta, 156:183-207. doi:10.1007/s00604-006-0623-9.

Seisonen, S, Vene, K & Kopple, K. 2016. The current practice in the application of chemometrics for correlation of sensory and gas chromatographic data. Food chemistry, 210:530-540. doi:10.1016/j.foodchem.2016.04.134.

Sell, CS. 2006. On the unpredictability of odor. Angewandte chemie international edition, 45:6254-6261. doi:10.1002/anie.200600782.

Sezille, C, Fournel, A, Rouby, C, Rinck, F & Bensafi, M. 2014. Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists, and perfumers. Frontiers in psychology, 5(12). doi:10.3389/fpsyg.2014.00012.

Shang, L, Liu, C, Tomiura, Y & Hayashi, K. 2017. Machine-learning-based olfactometer: Prediction of odor perception from physicochemical features of odorant molecules. Analytical chemistry, 89(22):11999-12005. doi:10.1021/acs.analchem.7b02389.

Shanteau, J. 1992. Competence in experts: The role of task characteristics. Organizational behavior and human decision processes, 53(2):253-266. doi:10.1016/0749-5978(92)90064-E.

Sicard, M, Baudrit, C, Leclerc-Perlat, M-N, Wuillemin & Perrot, N. 2011. Expert knowledge integration to model complex food processes. Application on the camembert cheese ripening process. Experts systems with applications, 38(9):11804-11812. doi:10.1016/j.eswa.2011.03.068.

Small, DM. 2012. Flavor is in the brain. Physiology and behavior, 107(4):540-552. doi:10.1016/j.physbeh.2012.04.011.

Small, DM, Gerber, JC, Mak, YE & Hummel, T. 2005. Differential neural responses evoked by orthonasal versus retronasal odorant perception in humans. Neuron, 47(4):593-605. doi:10.1016/j.neuron.2005.07.022.

Small, DM & Prescott, J. 2005. Odor/taste integration and the perception of flavor. Experimental brain research, 166:345-357. doi:10.1007/s00221-005-2376-9.

Snitz, K. 2016. Describing odorspace. P63. In: XXVIth Annual meeting of the european chemoreception research organization, ECRO 2016: Athens, Greece, 7-10 September, 2016, Organizer: Marika Kapsimali. doi:10.1093/chemse/bjw120.

Snitz, K, Yablonka, A, Weiss, T, Frumin, I, Khan, RM & Sobel, N. 2013. Predicting odor perceptual similarity from odor structure. Plos computational biology, 9(9):e1003184. doi:10.1371/journal.pcbi.1003184.

Solomon GEA. 1997. Conceptual change and wine expertise. The journal of the learning sciences, 6(1):41-60. doi:10.1207/s15327809jls0601_3.

Spehr, M, Gisselmann, G, Poplawski, A, Riffell, JA, Wetzel, CH, Zimmer, RK & Hatt, H. 2003.Identification of a testicular odorant receptor mediating human sperm chemotaxis. Science, 299(5615):2054-2058. doi:10.1126/science.1080376.

Stone, H & Sidel, JL. 2004. Sensory evaluation practices: Third edition, ed Taylor, SL. Elsevier academic press.

Strigler, F, Touraille, C, Sauvageot, F, Barthélémy, J, Issanchou, S & Pagès, J. 2009. Les épreuves discriminatives et descriptives. In: Evaluation sensorielle - Manuel méthodologique, eds Depledt, F & SSHA. Editions Tec & Doc, Lavoisier, Paris, France. 99-194.

Szulczyński, B, Gębicki, J & Namieśnik, J. 2018. Application of fuzzy logic to determine the odour intensity of model gas mixtures using electronic nose. E3S web of conferences, 28:01036. doi:10.1051/e3sconf/20182801036.

T - U

Tempere, S, Hamtat, M-L, de Revel, G & Sicard, G. 2016. Comparison of the ability of wine experts and novices to identify odorant signals: A new insight in wine expertise. Australian journal of grape and wine research, 22(2):190-196. doi:10.1111/ajgw.12192.

Thomas-Danguin, T. 2009. Flavor. In: Encyclopedia of Neuroscience, eds Binder, MD, Hirokawa, N & Windhorst, U. Berlin: Springer-Verlag Berlin and Heidelberg GmbH & Co. K. 1580-1582. doi:10.1007/978-3-540-29678-2_1772.

Thomas-Danguin, T & Chastrette, M. 2002. Odour intensity of binary mixtures of odorous compounds. Comptes rendus biologies, 325(7):767-772. doi:10.1016/S1631-0691(02)01485-3.

Thomas-Danguin, T, Maîtrepierre, E, Sigoillot, M, Briand, L & Tromelin, A. 2012. Les molécules odorantes, sapides et trigéminales. In: Odorat et goût : De la neurobiologie des sens chimiques aux applications, eds Salesse, R & Gervais, R. Editions Quae, Versailles, France. 29-37.

Thomas-Danguin, T, Sinding, C, Romagny, S, El Mountassir, F, Atanasova, B, Le Berre, E, Le Bon, A-M & Coureaud, G. 2014. The perception of odor objects in everyday life: A review on the processing of odor mixtures. Frontiers in psychology, 5:504. doi:10.3389/fpsyg.2014.00504.

Tomasino, E & Tomasino AP. 2017. Applying fuzzy-set logic analysis to relationships between flavor chemistry and sensory perception: A case of red fruit aromas in wine. 254th American Chemical Society National Meeting, Washington D.C., USA, 20-24 August.

Tromelin, A, Andriot, I & Guichard, E. 2006. Protein-Flavour interactions. In: Flavour in food, eds Voilley, A & Etiévant, P. Woodhead publishing. 172-207.

Tromelin, A, Chabanet, C, Audouze, K, Koensgen, F & Guichard, E. 2017. Multivariate statistical analysis of a large odorants database aimed at revealing similarities and links between odorants and odors. Flavour and fragrance journal, 33:106-126. doi:10.1002/ffj.3430.

Ullrich, F & Grosch, W. 1987. Identification of the most intense odor compounds formed during autoxidation of linoleic acid. Zeitschrift für Lebensmittel-Untersuchung und -Forschung, 184(4):277-282. doi:10.1007/BF01027663.

Upadhyay, R, Sehwag, S & Mishra, HN. 2017. Electronic nose guided determination of frying disposal time of sunflower oil using fuzzy logic analysis. Food chemistry, 221:379-385. doi:10.1016/j.foodchem.2016.10.089.

V - W - X - Y

van Ruth. 2001. Methods for gas chromatography-olfactometry: A review. Biomolecular engineering, 17:121-128. doi:10.1016/S1389-0344(01)00070-3.

van Ruth, SM, King, C & Giannouli, P. 2002. Influence of lipid fraction, emulsifier fraction, and mean particle diameter of oil-in-water emulsions on the release of 20 aroma compounds. Journal of agricultural and food chemistry, 50(8):2365-2371. doi:10.1021/jf011072s.

Vassar, R, Ngai, J & Axel, R. 1994. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell, 74(2):309-318. doi:10.1016/0092-8674(93)90422-M.

Vassiliadou, H & Lammert, M. 2011. Odeurs et dimension hédonique à travers le prisme des adjectifs. Langages, 181(1):73-88. doi:10.3917/lang.181.0073.

Veinand, B. 2015. Sensory testing with flavourists: Challenges and solutions. In: Rapid sensory profiling techniques and related methods: Applications in new product development and consumer research, eds Delarue, J, Lawlor, JB, Rogeaux, M. Woodhead publishing. doi:10.1533/9781782422587.3.383. 383-399.

Villière, A, Symoneaux, R, Roche, A, Eslami, A, Perrot, N, Le Fur, Y, Prost, C, Courcoux, P, Vigneau, E, Thomas-Danguin, T & Guérin, L. 2018. Dataset on the characterization of the flavor of two red wine varieties using sensory descriptive analysis, volatile organic compounds quantitative analysis by GC-MS and odorant composition by GC-MS-O. Zenodo [dataset]. doi:10.5281/zenodo.1213610.

Wise, PM, Olsson, MJ & Cain, WS. 2000. Quantification of odor quality. Chemical Senses, 25(4):429-443. doi:10.1093/chemse/25.4.429.

Xiao, Z, Wu, Q, Niu, Y, Wu, M, Zhu, J, Zhou, X, Chen, X, Wang, H, Li, J & Kong, J. 2017. Characterization of the key aroma compounds in five varieties of mandarins by gas chromatography−olfactometry, odor activity values, aroma recombination, and omission analysis. Journal of agricultural and food chemistry, 65(38):8392-8401. doi:10.1021/acs.jafc.7b02703.

Yan, L, Liu, J & Fang, D. 2015. Use of a modified vector model for odor intensity prediction of odorant mixtures. Sensors, 15(3):5697-5709. doi:10.3390/s150305697.

Yea, B, Konishi, R, Osaki, T & Sugahara, K. 1994. The discrimination of many kinds of odor species using fuzzy reasoning and neural networks. Sensors and actuators A: Physical, 45(2):159-165. doi:10.1016/0924-4247(94)00831-0.

Yu, P, Low, MY & Zhou, W. 2018. Design of experiments and regression modelling in food flavour and sensory analysis: A review. Trends in food science and technology, 71:202-215. doi:10.1016/j.tifs.2017.11.013.

Z

Zadeh, LA. 1965. Fuzzy sets. Information and control, 8(3):338-353. doi:10.1016/S0019-9958(65)90241-X.

Zareiforoush, H, Minaei, S, Alizadeh, MR & Banakar, A. 2015. A hybrid intelligent approach based on computer vision and fuzzy logic for quality measurement of milled rice. Measurement, 66:26-34. doi:10.1016/j.measurement.2015.01.022.

Zarzo, M. 2011. Hedonic judgments of chemical compounds are correlated with molecular size. Sensors, 11(4):3667-3686. doi:10.3390/s110403667.

Zarzo, M & Stanton, DT. 2006. Identification of latent variables in a semantic odor profile database using principal component analysis. Chemical senses, 31(8):713-724. doi:10.1093/chemse/bjl013.

Zhao, P, Gao, J, Qian, M & Li, H. 2017. Ccharacterization of the key aroma compounds in Chinese Syrah wine by gas chromatography-olfactometry-mass spectrometry and aroma reconstitution studies. Molecules, 22(7):1045. doi:10.3390/molecules22071045.

Zheng, Y, Sun, B, Zhao, M, Zheng, F, Huang, M, Sun, J, Sun, X & Li, H. 2016. Characterization of the key odorants in Chinese zhima aroma-type baijiu by gas chromatography-olfactometry, quantitative measurements, aroma recombination, and omission studies. Journal of agricultural and food chemistry, 64(26):5367-5374. doi:10.1021/acs.jafc.6b01390.

Zozulya, S, Echeverri, F & Nguyen, T. 2001. The human olfactory receptor repertoire. Biology, 2(6):research0018.1–research0018.12.

ANNEX

I. Publications

Publications appeared, submitted and in preparation (starting from the most recent).

- Arising from the thesis

Roche, A, Thomas-Danguin, T, Lutton, E & Perrot, N. 2018. Predicting the odor profile of food from its chemical composition with an artificial intelligence modelling approach coupling fuzzy logic and expertise. In writing.

Roche, A, Perrot, N & Thomas-Danguin, T. 2018. OOPS, the Ontology for Odor Perceptual Space: from molecular composition to sensory attributes of odor objects. In writing.

Roche, A, Thomas-Danguin, T, Perrot, N & Mailand, J. 2018. Predicting odor similarity of complex mixtures from molecular structure. In writing.

Villière, A, Symoneaux, R, Roche, A, Eslami, A, Perrot, N, Le Fur, Y, Prost, C, Courcoux, P, Vigneau, E, Thomas-Danguin, T & Guérin, L. 2018. Characterization of two red wine varieties using sensory descriptive analysis, volatile organic compounds (VOC) quantitative analysis and gas chromatography – olfactometry / mass spectrometry (GC-O-MS). Submitted to Data in Brief, Elsevier.

Roche, A, Perrot, N, Chabin, T, Villière, A, Symoneaux, R & Thomas-Danguin, T. 2017. In silico modelling to predict the odor profile of food from its molecular composition using experts' knowledge, fuzzy logic and optimization: Application on wines. 2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Montréal, Canada, 28-31 May. doi:10.1109/ISOEN.2017.7968875.

- Anterior to the thesis

Acree, TE, Roche, A, Charrier, A & Lavin, EH. 2015. Perception of odour mixtures: elements and images. Flavour Science, proceedings of the XIV Weurman Flavour Research Symposium. Context Products Ltd., Licestershire, UK, 279 - 284.

II. Communications arising from the thesis

Communications in French and international conventions (starting from the most recent).

European Chemoreception Research Organization (ECRO) annual meeting. Place: Würzburg, Germany. Date: 5-8 September 2018. Presentation: Poster titled "Predicting odor similarity of complex mixtures from molecular approach". Travel grant ECRO 2018 award obtained from the European Chemoreception Research Organization.

Human Chemosensation 2018 - the lab meeting. Place: Dresden, Germany. Date: 22 - 24 February 2018. Presentation: Poster titled "How to predict the odor profile of complex odor mixtures from their chemical composition?".

Journée des doctorants (PhD students' day). Place: INRA Dijon, France. Date: 19 December 2017. Presentation: 15 minutes talk titled "Wine Tuesday: *Prédiction de l'odeur de vins [2 modèles pour le prix d'un]*".

Pangborn 2017, Sensory Science Symposium. Place: Providence, Rhode Island, USA. Date: 20 - 24 August 2017. Presentation: Poster titled "Predicting the aromatic profile of wines from their chemical composition: Similarity among wines at different levels".

ISOEN 2017, The International Symposium on Olfaction and Electronic Nose. Place: Montréal, QC, Canada. Date: 28 - 31 May 2017. Presentation: Poster titled "How to predict the odor profile of wine from its chemical composition? In silico modelling using experts' knowledge, fuzzy logic and optimization".

EuroSense 2016, Seventh European Conference on Sensory and Consumer Research. Place: Dijon, France. Date: 11 - 14 September 2016. Presentation: Poster titled "In silico modelling to predict the odor profile of red wines from their molecular composition using experts' knowledge, fuzzy logic and optimization".

Journée des doctorants (PhD students' day). Place: INRA Dijon, France. Date: 12 August 2016. Presentation: 15 minutes talk titled "*Vini, Senti, Prédit*: How to predict the odor of a wine?".

Forum des Jeunes Chercheurs (Young researchers' event). Place: Université de Besançon, France. Date: 16 June 2016. Presentation: Poster titled "In silico modelling to predict the odor

profile of food from its molecular composition using experts' knowledge, fuzzy logic and optimization".

III. Award

Travel grant ECRO 2018 award obtained from the European Chemoreception Research Organization (ECRO) to attend the ECRO annual meeting in Würzburg (Germany).

IV. Scientific popularization events

Scientific popularization events in French.

Salon International de l'Agriculture (Paris International Agricultural Show). Place: Paris, France. Date: March 2016. Activity: Teaching visitors about the 5 tastes (Sweet, Salty, Bitter, Sour and Umami) with sensory tastings and questionnaires.

V. International environment

In the context of my PhD work, I was selected to be part of the doctoral program of the Agreenium International School of Research (EIR-A). This program aims to enhance the knowledge of PhD students with large scientific questions during two seminars (the first one was about "Sustainable and efficient food system: food, water and energy nexus", the second one will be about "Livestock issues of tomorrow's society"). Moreover this program is promoting the international employability of PhD students by encouraging them to do a scientific internship abroad.

I spend five months abroad, from June 1st 2017 to November 1st 2017, as a Visiting Scholar at the Monell Chemical Senses Center (Philadelphia, USA). This is the world leading center for the study of taste and smell. I was under the supervision of Dr. Joel Mainland. From this

collaboration, a publication is in writing (Chapter 2) and further scientific questions might be addressed as future collaborations.

VI. Training program and seminars

In addition to my thesis project, I had the opportunity to follow university courses and to participate at seminars.

Concerning the courses I studied mainly statistics (R) and modelling. Concerning the modelling, I followed courses about MATLAB programing and a two-week course titled "When nature inspires engineers: evolutionary algorithms, fuzzy logic, artificial neurons" at AgroParisTech (Paris).

I also studied English and attended first-aid training. I obtained my workplace first-aider diploma in 2017.

Finally, I attended a scientific seminar titled "Integration of data, knowledge and models: from the organization of information to its modelling" within the INRA facility.

ABSTRACT

Among the sensory dimensions involved in food flavor, the odor component is critical because it often determines the identity and the typicality of the food. Chemical flavor analysis provides a list of the odorants contained in a food product but is not sufficient to predict the odor resulting from their mixture. Indeed, odor perception relies on the processing by the olfactory system of many odorants embedded in complex mixtures and several perceptual interactions can occur. Thus, the prediction of the perceptual outcome of a complex odor mixture remains challenging and two main approaches emerge from the literature review. On the one hand, predictive approaches based on the molecular structure of odorants have been proposed but have been limited to single odorants only. On the other hand, methodologies relying on recombination strategies after the chemical analyses of flavor have been successfully applied to identify those odorants that are key to the food odor. However, the choices of odorants to be recombined are mostly based on empirical approaches. Thus, two questions arise: How can we predict the odor quality of a mixture on the basis of the molecular structure of its odorants? How can we improve food flavor analysis in order to predict the odor of a food containing several tens of odorants? These two questions are at the basis of the thesis and of this manuscript which is divided in two main axes.

The first axis describes the development of a model based on the concept of angle distances computed from the molecular structure of odorants in order to predict the odor similarity between mixtures. The results highlight the importance of taking into account the odor intensity dimension to reach a good prediction level. Moreover, several perspectives are proposed to extend the model prediction beyond the similarity dimension and to predict more qualitative dimensions of odors.

The second axis presents an innovative strategy which allows integrating experts' knowledge in the flavor analysis procedure. Three different types of heterogeneous data are embedded in a mathematical model: chemical data, sensory data and knowledge from expert flavorists. Experts' knowledge is integrated owing to the development of an ontology, which is further used to define fuzzy rules optimized by evolutionary algorithms. The final output of the model is the prediction of red wines' odor profile on the basis of their odorants' composition. Overall, the thesis work brings original results allowing a better understanding of food odor construction and gives insights on the underlying relationships within the odor perceptual space for complex mixtures.

Keywords: Food flavor, Olfaction, Mixture of odorants, Predictive modelling, Expert knowledge, Fuzzy logic.

RESUME

Parmi les dimensions sensorielles engagées dans la perception de la flaveur, la composante odorante est déterminante car elle porte le plus souvent l'identité et la typicité d'un aliment. L'analyse chimique de la composante odorante repose sur une stratégie séparative qui permet d'identifier les différents odorants présents dans l'aliment. Cependant, la perception des odorants en mélange induit des interactions au niveau perceptif qui ne sont pas prises en compte dans les techniques séparatives. Les mécanismes sous-jacents aux interactions perceptives sont mal connus, ce qui limite les possibilités de prédiction de l'odeur d'un aliment sur la base de sa composition chimique. En réponse à cette problématique deux approches émergent de la revue de la littérature. La première est basée sur la prédiction d'odeur d'après la structure moléculaire des odorants. Cependant, les études concernent des odorants seuls et non leurs mélanges. La seconde repose sur la recombinaison d'odorants en mélange après l'étape d'analyse séparative, mais le choix des odorants à associer est essentiellement empirique. Ainsi, deux questions se posent : Comment prédire l'odeur de mélanges de molécules d'après la structure moléculaire des odorants? Comment améliorer l'analyse de la flaveur dans le but de prédire l'odeur d'aliments complexes composés de plusieurs dizaine d'odorants en mélanges? Ces deux questions ont été abordées dans cette thèse dont les travaux sont décrits dans ce manuscrit selon deux axes principaux.

Le premier axe décrit l'utilisation et le développement d'un modèle basé sur le concept des distances angulaires calculées à partir de la structure moléculaire des odorants avec pour objectif de prédire la similarité perceptive de mélanges plus ou moins complexes d'odorants. Les résultats soulignent l'importance de prendre en compte la dimension d'intensité des odorants afin d'améliorer la qualité de la prédiction. Des perspectives d'amélioration du modèle sont dégagées pour permettre de dépasser la dimension de similarité et prédire des dimensions qualitatives de l'odeur.

Le deuxième axe présente une démarche originale d'intégration de connaissances liées à l'expertise dans la procédure d'analyse de la flaveur. Ainsi, trois types de données hétérogènes sont agrégés dans un modèle mathématique global : des données chimiques, des données sensorielles et des connaissances d'experts aromaticiens. L'expertise est intégrée à travers la création d'une ontologie qui est ensuite associée à une approche de logique floue optimisée par algorithme évolutionnaire. Le modèle développé permet de prédire le profil odorant de seize vins rouges sur la base de leur composition en odorants. Au final, l'ensemble des travaux menés dans cette thèse apporte des résultats originaux permettant une meilleure compréhension de la construction des odeurs des aliments et permet d'élaborer des hypothèses quant aux relations sous-jacentes de l'espace perceptif des odeurs en mélanges complexes.

Mots-clés : Flaveur des aliments, Olfaction, Mélange d'odorants, Modélisation prédictive, Expertise, Logique floue.