Odor naming is difficult, even for wine and coffee experts

Ilja Croijmans¹
(i.croijmans@let.ru.nl)

Asifa Majid¹,²,³
(asifa.majid@let.ru.nl)

¹Centre for Language Studies, Radboud University, Nijmegen, NL
²Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, NL
³Max Planck Institute for Psycholinguistics, Nijmegen, NL

Abstract

Odor naming is difficult for people, but recent cross-cultural research suggests this difficulty is culture-specific. Jahai speakers (hunter-gatherers from the Malay Peninsula) name odors as consistently as colors, and much better than English speakers (Majid & Burenhult, 2014). In Jahai the linguistic advantage for smells correlates with a cultural interest in odors. Here we ask whether sub-cultures in the West with odor expertise also show superior odor naming. We tested wine and coffee experts (who have specialized odor training) in an odor naming task. Both wine and coffee experts were no more accurate or consistent than novices when naming odors. Although there were small differences in naming strategies, experts and non-experts alike relied overwhelmingly on source-based descriptions. So the specific language experts speak continues to constrain their ability to express odors. This suggests expertise alone is not sufficient to overcome the limits of language in the domain of smell.

Keywords: Olfaction; Flavor Expertise; Odor Naming; Language Consistency; Cognitive Linguistics

Introduction

Strawberries, stop signs and fire trucks are red. It is easy to name the common color property “red” irrespective of the object it belongs to. In fact, many (if not all) languages have dedicated or “basic” color terms to refer to hues regardless of the object to which they belong (Berlin & Kay, 1969; Kay, Berlin, Maffi, & Merrifield, 2009). That is, there is a set of abstract terms which only refer to the quality of color (e.g., blue, green, red). Color appears to be highly “codable”, i.e., easy to express in language. Brown and Lenneberg (1954) operationalized codability with a number of parameters, aside from basic vocabulary. Descriptions for codable colors are shorter (measured in number of syllables, and number of words); are named faster; and people agree on the names (there is consensus both across people and within people across time).

In contrast to colors, odors do not appear to be codable. Philosophers and scholars throughout history argue smell is the least important sense (e.g. Darwin, 1871; Sperber, 1975/1974), and this appears to be reflected in language too. People are generally very bad at naming smells. Ask two people to name the same odor and they produce different descriptions; in most cases referring to a source (e.g., fruit, banana) or alternatively giving an evaluative response (e.g., gross, beautiful) (Lawless & Cain, 1975; Lawless, 1984; Yeshurun & Sobel, 2010). There appear to be no basic words for smell (Sperber, 1975/1974). Odors appear to be ineffable (cf. Levinson & Majid, 2014).

However, a recent study shows poor odor naming might be a product of WEIRD (Western, Educated, Industrialized, Rich and Democratic) cultures (Henrich, Heine, & Norenzayan, 2010). Majid and Burenhult (2014) tested Jahai speakers, a group of hunter-gatherers living on the Malay Peninsula in an odor naming task, and compared them to a matched sample of English speakers. The Jahai were as consistent in naming odors as they were in naming colors. Moreover, they were more consistent in naming odors than English speakers. A more qualitative look at the data showed that while English speakers overwhelmingly used source-based descriptions to name smells (as found in previous studies), Jahai speakers used a small set of abstract or “basic” smell terms. These are words which exclusively refer to smell qualities: the terms do not derive from a source; they apply to a broad class of objects; they are psychologically salient to the Jahai; and they appear in all genres of conversation (Burenhult & Majid, 2011). There are as many as 12 basic odor terms in Jahai. For example, the smell of bat droppings, smoke, ginger root, and petroleum are all described with the word cyes, and the smell of various flowers, perfumes, durian, and bearcat (Arctictis binturong) are named lipit.

For the Jahai smell plays a significant role, not only in language, but in various other facets of life, such as religion and medicine (Burenhult & Majid, 2011). For example, it is taboo to wash raw game meat from different species at the same time in the river. This is because the blood of each animal has a distinct smell and mixing these smells is prohibited. For the Jahai a cultural preoccupation with odors aligns with high codability of smells in language. This raises the question of whether other cultural practices involving smell may also lead to greater ease in odor naming.

In the West, smells play a significant role in the lives of various professionals, for example, perfumers and flavor experts (such as wine and coffee experts). Flavor is a multisensory percept, consisting of orthonasal (through the nose), and retronasal (through the mouth) olfaction alongside taste (for a review on the multisensory perception of flavour see Auvray & Spence, 2008). Thus, olfaction is in large part responsible for how we perceive the taste of the things we eat and drink. Wine, coffee, cheese, and chocolate would all taste bland without the sense of smell. For this reason a large part of flavor experts’ training focuses on olfaction (cf. Herdenstam, Hammarên, Ahlstrôm, &
Wiktorsson, 2009). Communicating about smells is also part of the sub-culture of flavor expertise. Think for example of a vinologist who explains to a customer why this particular Bordeaux vintage, rather than the Beaujolais primeur the customer had in mind, is better suited to a lamb stew. So the question we ask is whether flavor experts in the West are also good odor namers. Could the specialized training flavor experts undergo make them better at naming smells than the lay person?

The use of language by experts has received little attention in this context. There are two separate issues: (1) Are experts more consistent than non-experts in describing smells? (2) Do experts differ from non-experts in the types of strategy they use to describe smells? Previous studies have not directly examined whether everyday odor naming is different between experts and novices. The studies suggest a rather mixed picture of experts’ linguistic abilities for odors.

Some studies suggest experts might not be much better than novices at odor naming. Lawless (1984) asked wine experts and novices to describe different wines, and looked at the content of their descriptions. He found wine descriptions from both experts and novices were highly idiosyncratic suggesting little systematicity in expert responses. A whopping 82% of odor descriptions were given by only one participant for each wine, indicating low consistency across participants. However, Lawless did not directly compare the two groups on consistency, so we cannot know for sure whether experts and novices were similar or different on this measure. In a different study Parr, Heatherbell, and White (2002) asked wine experts and matched novices to sniff and then name wine related odors (instead of actual wine). Experts and novices correctly identified the same number of odors, and were equally consistent. However, a closer look at the data shows experts had numerically higher identification and consistency rates (even though these did not prove statistically different), leaving open the possibility the study was underpowered (as suggested by Parr et al. [2002, p.752] themselves).

In contrast to Lawless and Parr et al., Bende and Nordin (1997) found experts named more odors correctly than novices, suggesting experts were also more consistent. The expert advantage has been found in director-matcher tasks too. In this paradigm people are asked to match wines to linguistic descriptions produced by experts and novices. Expert descriptions led to more correct matches than novice descriptions for the same wines (Solomon, 1990). This suggests expert descriptions are more informative. Experts are also better than novices at matching descriptions to wines, particularly when the descriptions are given by other experts (Lawless, 1984). Taken together these results suggest flavor experts may indeed have an advantage for naming odors.

In terms of the strategies experts use, some studies have found wine experts use more concrete and specific words (e.g. blackberries) (Lawless, 1984), or more precise language overall (Zucco, Carassai, Baroni, & Stevenson, 2011). In contrast, when matching descriptions to wine, Gawel (1997) found wine experts relied more on vague or metaphorical terms (e.g. elegant, complex). In another study, Sezille, Fournel, Rouby, Rinck, and Bensaï (2014) compared naming in experts (perfumers, flavorists) and non-experts (novices, trainee chefs). They found experts used more technical terms (i.e., terms referring to chemical substances (e.g. aldehyde), and made less reference to the hedonic value (e.g. unpleasant).

The studies reviewed above focus mostly on odors relevant to the specific expertise being considered, so it is not clear if flavor expertise leads to a general advantage in naming odors. Do the many years of experience change wine experts’ ability to express odors in language? In this study we compared a group of wine experts to non-experts for their naming of simple everyday odors (such as chocolate, leather, and lemon). We also included a separate group of coffee experts. Coffee experts also undergo extensive training of their noses and palates, but differ in some interesting respects from wine experts. Whereas wines are usually elaborately described in tasting notes, menus, and on placards in stores, the descriptions of coffees tend to be simpler. This is illustrated by the number of magazines available to peruse on both topics; while there are 10 different subscription magazines to be found about wine on Amazon.com, not a single subscription magazine for coffee is available (retrieved on January 20th 2015). This suggests wine experts have more experience in communicating about odors than coffee experts, even though both have extensive perceptual experience and training with smells and flavors.

We test three main hypotheses in this study. First, given experts’ greater training and everyday attention to smells, experts should be more consistent than novices in the descriptions of smells. This might be especially true for wine experts who have many opportunities to communicate about smells. Second, experts should also be able to correctly identify more smells than novices. Finally, experts should differ in the precision of their descriptions for odors. Based on the previous literature it is unclear what the direction of this difference should be exactly, but we might expect experts to differ both in the length of their descriptions and the type of descriptions they give.

**Methods**

**Participants**

Sixty-three people (22 women, $M_{age} = 43.7$ years, $SD = 11.7$, age range: $24 – 70$ years) participated in the experiment. Twenty-two participants were wine experts, and worked as qualified vinologists or sommeliers. Twenty participants were coffee experts, and worked as qualified baristas, coffee roasters or coffee brokers, and had a similar amount of training and experience as the wine experts. Another 21 participants were recruited as novice controls and were matched for age and gender to both the wine and coffee expert groups as closely as possible. All participants were native or near-native speakers of Dutch.
To assess expertise, all participants completed a questionnaire which tested their knowledge about wine (following Hughson & Boakes, 2001; Lehrer, 1983), their knowledge about coffee,¹ and their general awareness of odors in daily life (a shortened version of the questionnaire by Smeets, Schifferstein, Boelma, & Lensvelt-Mulders, 2008). Separate ANOVAs on the three questionnaires confirmed the wine experts have expertise in wine (and only wine), the coffee experts in coffee (and only coffee), and the novices in neither. Both wine and coffee experts showed significantly higher odor awareness than the novices, but they did not differ from each other in this respect.

Materials
Ten different odors were used in the odor naming task presented using “Sniffin’ Sticks” (Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997). These are marker pens containing an odorant (instead of ink) which the participant can smell by removing the marker cap and smelling the tip. All odors used in the experiment were common household scents, and were familiar to people living in the Netherlands. The odors can also be found in the descriptions of wine and coffee (e.g. Noble et al., 1984). The odors were: chocolate, clove, apple, lemon, cinnamon, garlic, mushroom, leather, grass, rose.

Procedure
Each participant was tested individually in a neutral, well-lit and well-ventilated room, kept at 20±2 degrees Celsius. The entire procedure was carried out in Dutch. The participants completed the questionnaires first, and then the odors were presented in a fixed random order. Each participant was instructed in Dutch: Wilt u de geur zo precies mogelijk beschrijven (“Can you describe the smell as precisely as possible”). Answers were recorded using an audio-recorder, and the descriptions were later transcribed and coded.

Coding
First we measured the length of every description by counting the number of characters in the fully transcribed response.

We wished to examine whether experts and novices agreed in their responses, and whether they correctly identified the odors. In order to measure this, the main responses from the fully transcribed descriptions were identified. For example, a speaker said: Ruikt wel naar chocola. Vanille of chocola een van de twee. Geen idee. i.e., ‘Smells like chocolate. Vanilla or chocolate, one of the two. No idea.’ From this description the main qualitative descriptors chocola and vanille were coded. Modifiers and hedges were ignored unless their exclusion changed the quality description. For example, een beetje ‘a little’ in een beetje citrus ‘a little citrus’ was not coded because the quality is “citrus”, and een beetje indicates only the strength of the odor (or confidence of the participant). But rood fruit ‘red fruit’ was coded as a whole response including rood ‘red’, because “red fruit” has a different quality of smell than “fruit” in general. Repeated responses (e.g. when a person mentioned chocola twice for the same stimulus as in the example above) were only counted once.

Consistency between speakers was calculated as in Majid and Burenhult (2014) using Simpson’s Diversity Index (Simpson, 1949), and was measured separately for only first responses, and then all responses. That is, for each stimulus item we calculated whether each participant in each group agreed with one another in how they described that item. Item scores were subjected to further analyses (see Results). For accuracy, answers were coded as correct when the participant identified the odor as classified by the manufacturer of the stimuli. The total number of correct responses was then calculated over participants and items.

Finally, we coded responses into three main categories so we could test whether experts differed from novices in the strategies they used: (1) Source terms, i.e. words referring to an object or class of objects (e.g., chocola ‘chocolate’, frutig ‘fruity’); (2) Evaluative terms, i.e. words describing the hedonic evaluation of the stimulus (e.g., lekker ‘pleasant’, gadverdamme ‘disgusting’). Majid and Burenhult (2014) identified a third category of abstract terms. In Dutch this includes terms such as aromatisch ‘fragrant/aromatic’ and muf ‘musty’. Participants rarely used this strategy but they did use other descriptions, such as cross-modal metaphors (e.g., zoet ‘sweet’, groen ‘green’), or reference to a general state (e.g., gekookt ‘cooked’). We categorized these together with abstract odor terms as (3) Non-source terms, i.e. words not referring directly to an object.

Results
We first examined the precision of responses across groups by comparing the length of descriptions using one-way ANOVAs (3 levels: wine expert, coffee expert, novices) by participants (F₁) and items (F₂). There was a significant main effect of group F₁ (2, 56) = 11.8, p < 0.001, η² = 0.037, F₂ (2, 27) = 5.812, p = 0.008, η² = 0.30. Planned comparisons showed coffee experts (M = 102, SD =103) gave significantly shorter descriptions than wine experts (M = 146, SD = 125), p < 0.001, d = 0.38, and novices (M = 144, SD =127), p < 0.001, d = 0.36. The difference between wine experts and novices was not significant. So, coffee experts were more concise in their descriptions overall.

Our main question was whether experts are more consistent when naming odors. A one-way ANOVA on consistency scores across items showed no main effect of group by first response F (2, 27) = 0.904, p = 0.417, η² = 0.063, or all responses F (2, 27) = 1.251, p = 0.302, η² =

¹ A coffee questionnaire was specifically designed for this study. It was modelled on previous wine questionnaires, and was developed with the assistance of Rose van Asten, a qualified Specialty Coffee Association Europe (SCAE) coffee expert.
Experts were no more consistent for odor descriptions than novices. Scores were in the same range for all groups, and relatively low throughout (ranging from 0.07 to 0.12 where the maximum score indicating unanimity would be 1.0; See Figure 1).

Figure 1: The bar graph (top) illustrates overall consistency (for first responses) was low for all groups and not statistically different from each other (error bars represent standard error). The pie charts (below) illustrate types of descriptions. All three groups predominantly used source terms (orange). In addition, wine experts used more non-source terms (black), and novices more evaluative terms (green).

We then examined whether experts correctly identified the target odors more often by conducting one-way ANOVAs again by participants ($F_1$) and items ($F_2$). There was no difference between groups in the number of correct responses $F_1 (2, 59) = 1.356$, $p = 0.266$, $\eta^2 = 0.044$, $F_2 (2, 27) = 0.094$, $p = 0.910$, $\eta^2 = 0.007$. Wine experts ($M = 5$, $SD = 2.3$) and coffee experts ($M = 6$, $SD = 4.2$) were no better than novices ($M = 5$, $SD = 4.0$) in identifying odors. Finally, we examined the type of descriptions across groups using chi-square. All groups overwhelmingly used source-based terms (wine experts 70%; coffee experts 81%; novices 72%), but there was an overall effect of strategy by group, $\chi^2 (4, N = 1698) = 22.9$, $p < 0.01$, Cramer’s $V = 0.074$. Wine experts used more non-source terms (e.g., zoet ‘sweet’, gekookt ‘cooked’), $z = 2.18$, $p < 0.05$, while novices used more evaluative terms (e.g. lekker, i.e. ‘nice’), $z = 1.94$, $p = 0.052$. Compared to the other groups, coffee experts used fewer non-source terms, $z = -1.978$, $p < 0.05$, and fewer evaluative terms, $z = -2.33$, $p < 0.05$ (see Figure 1). So the three groups differ in subtle ways in their overall strategy for describing smells, but the dominant strategy (i.e., reliance on source descriptions) is the same for all.

### Discussion

Wine and coffee experts were no more consistent than novices when describing odors. In fact, overall consistency for odor naming was low for both experts and novices, and comparable to the English speakers reported by Majid and Burenhult (2014). In addition, we found no difference between groups in the number of correctly labeled odors. Overall, experts and novices all overwhelmingly used source-based descriptions. There were, however, small differences in specific strategies. Coffee experts gave the shortest descriptions, and used less evaluative terms and non-source terms than wine experts or novices. Wine experts used more non-source, or metaphorical, terms to describe the odors, such as zoet ‘sweet’ and gekookt ‘cooked’, while novices used more evaluative descriptions, such as lekker ‘pleasant’ and gadverdamme ‘disgusting’.

The odors used in this study represented everyday smells. Nevertheless, experts and novices had problems accurately naming them. This is in line with the results of Parr and colleagues (2002), who found experts and novices do not differ in their ability to name smells. Our study had double the number of wine experts and novices as Parr et al., (2002) and included an additional 20 coffee experts, so the lack of an effect is unlikely due to insufficient power. In fact, the wine experts show numerically lower consistency than either coffee experts or novices. It is also unlikely the absence of an effect was due to lack of expertise. All experts in this study fulfilled the criteria for expertise (cf. Melcher & Schooler, 1996; Parr et al., 2002): they were wine or coffee professionals, earning money with their expertise. We also independently verified the experts tested really did have more knowledge of wines and coffees by means of Expert Questionnaires. The results of the odor awareness scale also showed experts differed from novices by showing more awareness of odors in general; further confirming odors play an important role for them in their daily lives.

Some previous studies have found an advantage of expertise on odor naming. Bende and Nordin (1997) found experts named more odors correctly than novices, but a closer examination of the data shows the group differences were driven by only a few odors in the sample. In a similar vein, Zucco and colleagues (2011) found wine experts were able to identify wine related odors with higher accuracy than novices. However, this advantage did not generalize to common household odors. Our results corroborate this finding. Expert training in flavor does not give a general advantage for naming smells. However, experts could very well be better at naming and identifying specific smells, particularly those relevant for their expertise. For example, wine experts could be better at identifying odors specifically

---

2 Separate participant analyses are not possible for the consistency analyses because consistency is measured over participants. The results were verified with the non-parametric Kruskall-Wallis test, as the assumptions for a parametric test were violated. The pattern of results is the same.
related to wine (e.g. smells related to barrel ageing) and coffee expertise for coffee-related smells (e.g. smells related to the Maillard reaction). Although the smells used in this study all occur in wine and coffee descriptions (e.g. Noble et al., 1984), they are found elsewhere too and do not require any special training. Another possibility is that remains open is experts are more proficient in naming smells and flavors of real substances from their domain of expertise, i.e. wines and coffees. This requires further testing.

Although there was no difference between the groups in consistency, the groups did differ in subtle ways in the types of responses they gave. The difference between the two expert groups is a little surprising given previous research. Sezille and colleagues (2014) asked flavorists, perfumers, trainee cooks and novices to describe different odors. They showed a general tendency for novices to use more evaluative terms (as we also found in this study), but no differences between expert groups. In our study, both expert groups reported the same level of awareness of odors, so it is unlikely the difference in naming resides there. Instead, we suggest the variation comes from differences in evaluating the flavor of wines versus coffees.

The differences in odor naming strategies in different flavor domains have not been compared directly, but there are some intriguing observations in the literature. Wine experts often describe wine in a structured way (cf. Lehrer, 1983), first evaluating the color (visually), then the aroma and bouquet by smelling the wine (i.e. orthonasal olfaction), and then the various taste aspects of the wine (i.e. gustation, mouthfeel, retronasal olfaction and finish). The terms used in these various steps, however, are not strictly defined (e.g. Brochet & Dubourdieu, 2001; Lehrer, 1983). Wine experts appear to integrate smell, texture, and taste into more synaesthetic language (cf. Caballero, 2007; Paradis & Eeg-Olofsson, 2013). This is evident in their use of more non-source terms for the smells in this experiment as well. When describing coffees, on the other hand, coffee experts describe smell with source descriptors, and taste with a small set of abstract taste terms. The perfect coffee has the right balance between sweet, bitter and sour (cf. Hayakawa et al., 2010). The present study suggests a closer look at the different strategies specific groups of flavor experts use to describe odors could be fruitful to explore further.

How well do the current findings mesh with the idea that the senses are differentially ineffable (Majid & Burenhult, 2014; Majid & Levinson, 2011)? According to Levinson and Majid (2014), expert interest in smell could overcome the relative ineffability of the sense of smell in a given language. In contrast to the language of the Jahai, Dutch (like English) has very few dedicated words for odors. The experts in this study had years of training and experience, but even with this experience, wine and coffee experts were restricted by the limitations of their language. Why is this so?

One possibility suggested by the earlier discussion is the specific expertise we investigated. Although flavor expertise relies on odor knowledge, flavor is far more than odor alone. Perhaps focusing on dedicated odor experts, such as perfumers or incense makers, would be better. A different possibility for the poor odor naming by wine and coffee experts is the delayed acquisition of their expertise. The Jahai learn their smell lexicon in the course of normal language acquisition: as children. The Dutch experts, on the other hand, only come to acquire their expertise in odors and flavors late in life and long after any critical period for language learning. It could be the late development of odor expertise puts restrictions on the ability to learn odor language by flavor experts. These possibilities require further exploration.

To conclude, the present study indicates the resources within a specific language can restrict the codability of a perceptual sense, and selective experience and training is not enough to overcome these restrictions. This suggests the specific language we speak puts constraints on how we communicate and make sense of the world.

Acknowledgments

This work was funded by The Netherlands Organization for Scientific Research: NWO VICI grant “Human olfaction at the intersection of language, culture and biology”. Thanks to all the participants involved in this study. Special thanks go to Michiel Buijs, Rose van Asten and Liesbeth Sleijster, to Nina Krijnen for advice and assistance, and to Laura Speed and Doris Richter genannt Kemmermann for comments on an earlier version of this paper.

References


Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, 130, 266–270.


488