

# English taste and smell adjectives as generalized emotional vocabulary

Bodo Winter (University of Birmingham, UK)

## Introduction

One of the primary functions of language is the expression of evaluation (e.g., Dam-Jensen & Zethsen, 2007; Hunston, 2011). Another important function is the expression of sensory perceptions (e.g., Miller & Johnson-Laird, 1976). These two domains—perception and evaluation—are intimately connected. It has been claimed that compared to relatively neutral visual words such as *yellow* and *purple*, taste and smell adjectives have an obligatory evaluative component (Levinson & Majid, 2014). Whereas visual adjectives may be highly positive or negative (e.g., compare *shiny* and *attractive* to *dark* and *ugly*), there also is a large number of very neutral terms. On the other hand, it is difficult to find taste and smell descriptors that are not highly valenced. For example, the smell descriptors *rancid*, *pungent* and *stinky* have strong negative connotations compared to the highly positive descriptors *fragrant* and *aromatic*.

In this paper, I will combine various perceptual and emotional norm datasets with corpora to look at the semantic prosody of English perceptual words, investigating their evaluative functions in context. Crucially, my aim is to show that taste and smell words form a “generalized emotional vocabulary” that can be used across domains, that is, outside of their primary perceptual contexts, such as the description of food or perfumes. In other words, I propose that the evaluative component of taste and smell adjectives is so strong that they are very frequently used metaphorically to describe non-taste and non-smell domains.

## Analysis of the Corpus of Contemporary American English

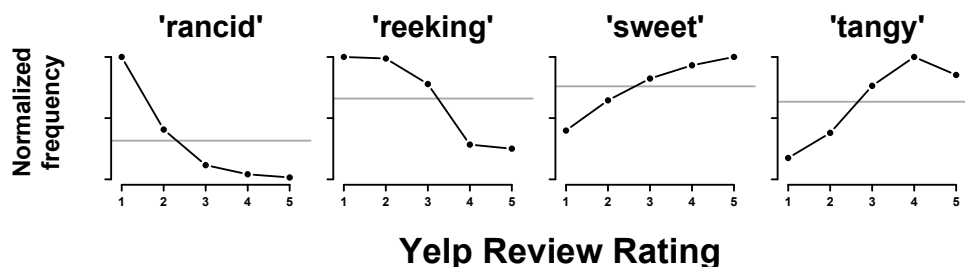
A set of sensory adjectives taken from Lynott and Connell (2009) will be analyzed. To construct this dataset, native speakers of English were asked to norm 400 words according to which perceptual modality they correspond to (e.g., *fragrant* is classified as “smell”, *purple* is classified as “visual”). These modality norms will be combined with emotional valence norms from Warriner, Kuperman and Brysbaert (2013), the SentiWordnet (Esuli & Sebastiani, 2006) and the Twitter Emotion Corpus norms (Mohammad & Kiritchenko, 2015). All of these datasets assign positive and negative numerical values to words according to their emotional valence. Crucially, the valence norms were constructed using different approaches: via human norms (Warriner et al., 2013), via a computational dictionary-based approach (SentiWordNet), and via a corpus-based approach based on occurrences with emotional terms (Mohammad & Kiritchenko, 2015).

In my first analysis, I use these emotional valence measures together with the Corpus of Contemporary American English (COCA, Davies 2008-), a large 520-million-word corpus balanced for five generic registers (newspaper, magazine, academic, literature, spoken). For the set of 400 adjectives, about 150,000 adjective-noun pairs (types) were extracted. A mixed linear regression model with perceptual modality as a fixed effect (sight vs. sound vs. touch vs. taste vs. smell) and word and noun as random effect (with by-modality random slopes) shows that across the entirety of COCA, taste adjectives overall modify more positive nouns and smell adjectives modify more negative nouns. By additionally using an absolute valence measure that quantifies overall emotional involvement (regardless of the sign of the emotional valence) it can furthermore be shown that taste and smell words overall occur with reliably more valenced nouns than words corresponding to

the other sensory modalities. For example, a visual word such as *yellow* frequently modifies relatively neutral nouns such as *trousers* or *house*. These results are obtained regardless for which emotional valence datasets are used.

## Analysis of online review data

The evaluative meaning of taste and smell words is most apparent for text types in which the main goal of speakers/writers is to express their subjective opinion. For this, online restaurant reviews from the Yelp Dataset Challenge will be used, comprising a set of 2.7 million UK and US restaurant reviews (of both U.K. and U.S. restaurants). Using the Yelp app, users can assign star ratings ("1 star", "2 stars" etc.). Potts (2011) has shown that this allows inferring the emotional meanings of words. Replicating the main findings above, taste words occurred overall in more positive reviews (high star rating) and smell words in more negative reviews (low star rating). Moreover, as the examples in Figure 1 show, review valence (in terms of star rating) yields monotonic increasing or decreasing functions for particular words.



**Figure 1.** Frequencies by rating (5 = most positive), cf. Potts (2011)

So far, the analyses have either aggregated across contexts (COCA) or investigated taste and smell words in food-related contexts (restaurant reviews). To truly show that taste and smell words characterize a domain-general set of affectively loaded terms that can be used across different topics, the current results need to be compared to descriptions for a non-food-related domain. Movie review data is excellent for this because movies are exclusively experienced through vision and sound. As the following sentence from Pang and Lee's (2004) online movie review dataset shows, movies are often talked about metaphorically in terms of food, using a high number of taste and smell words:

*Watching Trouble Every Day (...) is like biting into what looks like a juicy, delicious plum on a hot summer day and coming away with your mouth full of rotten pulp and living worms*

The perceptual words from Lynott and Connell (2009) were checked against the movie review corpus by Pang and Lee (2004), who categorized 10,000 sentences from the review site [www.rottentomatoes.com](http://www.rottentomatoes.com) for whether they were "objective" (occurring in the movie synopsis and describing the movie in a matter-of-fact style) or "subjective" (occurring in the evaluative part of the movie that expresses the critic's opinion). Analyses of these 10,000 sentences showed that taste and smell words were more likely to occur in the subjective sentences than in the objective sentences. Token frequencies were used to construct odds ratios (subjective versus objective). Taste, smell and touch words were more likely to occur in subjective than objective sentences (3.4, 2.0, 1.5), compared to sight and sound which showed a slight preference for objective sentences (0.7, 0.8). A Poisson regression model shows these differences to be reliable.

## Conclusions

Taken together, the findings support Levinson and Majid (2014), as well as Lehrer's (2009) analysis of perceptual adjectives in wine speech as performing both affective function as well as descriptive function. These results also call into question the proposal that expressions such as *sweet melody* are "synesthetic" (Shen, 1997; Yu, 2003; Shen & Aisenman, 2008), i.e., mappings between distinct sensory modalities (transfer of taste to sound). Instead of describing a *melody* as *sweet* in a truly gustatory sense, such expressions seem to be primarily involving the mapping of affective meanings. Altogether, the results suggest that taste and smell words form a subpart of the English lexicon that characterizes a domain-general emotional vocabulary.

## References

- Dam-Jensen, H., & Zethsen, K. K. (2007). Pragmatic patterns and the lexical system—A reassessment of evaluation in language. *Journal of Pragmatics*, *39*(9), 1608-1623.
- Davies, M. (2008-). *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>
- Esuli, A., & Sebastiani, F. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of 5th Conference on Language Resources and Evaluation* (Vol. 6, pp. 417-422).
- Hunston, S. (2011). *Corpus Approaches to Evaluation: phraseology and evaluative language*. London: Routledge.
- Lehrer, A. (2009). *Wine & Conversation*. Oxford: Oxford University Press.
- Levinson, S. C., & Majid, A. (2014). Differential ineffability and the senses. *Mind & Language*, *29*, 407-427.
- Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, *41*, 558-564.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge: Harvard University Press.
- Mohammad, S. M., & Kiritchenko, S. (2015). Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*, *31*, 301-326.
- Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistic (ACL)*.
- Potts, C. (2011). On the negativity of negation. In N. Li & D. Lutz (Eds.), *Proceedings of Semantics and Linguistic Theory*, *20* (pp. 636-659). Ithaca, NY: CLC Publications.
- Shen, Y. (1997). Cognitive constraints on poetic figures. *Cognitive Linguistics*, *8*, 33-71.
- Shen, Y., & Aisenman, R. (2008). Heard melodies are sweet, but those unheard are sweeter': Synaesthetic metaphors and cognition. *Language and Literature*, *17*(2), 107-121.
- Warriner, A.B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*, 1191-1207.
- Yu, N. (2003). Synesthetic metaphor: A cognitive perspective. *Journal of Literary Semantics*, *32*, 19-34.