Daniela Pettersson-Traba*

A diachronic perspective on near-synonymy: The concept of SWEET-SMELLING in American English

https://doi.org/10.1515/cllt-2018-0025

Abstract: This paper presents a diachronic analysis of the attributive uses of four synonymous adjectives which designate the concept of SWEET-SMELLING (*fragrant*, *perfumed*, *scented*, and *sweet-smelling*) in the latter part of Late Modern and Present-day American English. By drawing on data from the *Corpus of Historical American English* (*COHA*) and applying a Hierarchical Configural Frequency Analysis (HCFA), it delineates the internal semantic structure of this set of synonyms, paying special attention to their noun collocates. The results show that the concept of SWEET-SMELLING experiences major changes over the time span examined (1850–2009), from being used mostly to qualify entities which can exhibit a *natural* pleasant smell (e.g. flowers and trees) to modifying objects which are *artificially* sweet-smelling (e.g. oils and shampoos). Moreover, *fragrant* and *perfumed*, which initially were the most frequent adjectives, are gradually replaced by *scented*, thus reflecting a change in the relation between the synonyms over time. The study constitutes the first diachronic approximation to synonymy from the perspective of cognitive semantics and provides equally effective results as previous synchronic research in the field.

Keywords: synonymy, semantic change, collocational behavior, Hierarchical Configural Frequency Analysis

1 Introduction

Although absolute synonyms are virtually nonexistent, languages abound in near-synonyms, linguistic constructions with similar, though not identical meanings.[1] Near-synonyms are fundamental to the organization of the lexicon

1 Throughout this paper I employ the terms *synonym*/*synonymy* and *near-synonym*/*near-synonymy* interchangeably since many linguists concur on absolute synonymy being very infrequent in languages, if it exists at all.

***Corresponding author: Daniela Pettersson-Traba,** Department of English and German, University of Santiago de Compostela, Santiago de Compostela, Spain, E-mail: daniela.pettersson@usc.es
http://orcid.org/0000-0001-9302-3372

Published online October 09, 2018

of any language since they exert a great influence on the structure of lexical knowledge. For instance, synonyms lie at the basis of the creation and demarcation of meanings due to their *contrastive nature*. As claimed by Saussure (1983: 114),

> [i]n a given language, all the words which express neighbouring ideas help define one another's meaning. Each of a set of synonyms [...] has its particular value only because they stand in contrast with one another. [...] So, the value of any given word is determined by what other words there are in that particular area of the vocabulary.

To be considered synonymous, words or larger linguistic units must refer to the same concept in the extralinguistic reality, that is, they must share core denotational semantic traits. However, synonyms are permitted to contrast in peripheral aspects of their denotational meaning, or in other dimensions such as style, connotation, and collocation. Thus, synonyms typically convey different nuances and implications, making it difficult to choose the most suitable word from a group of near-synonyms in specific contexts of use.

Despite its complexity and importance, research on synonymy is scarce, especially if compared to other semantic phenomena such as polysemy and metaphor, which have been the subject of an ever-increasing body of literature (cf. Edmonds and Hirst 2002; Taylor 2003; Liu 2010). The semantic structure of specific groups of synonyms has received particularly little attention, and several researchers have recently stressed the need for more studies in this field. For instance, Divjak and Gries (2006) point out that "the issue of the internal structure of a set of synonyms has hitherto remained largely undiscussed in the literature", thus pointing to a clear existing gap in this area of research. Liu also draws attention to this fact in several of his studies on synonymy (Liu 2010, 2013; Liu and Espino 2012), claiming that "[a]dditional systematic research is needed to share more light on this intriguing linguistic phenomenon and, more importantly, to gain a better understanding of how various specific sets of synonyms work in terms of meaning and usage patterns" (Liu 2010: 57). This has led to the emergence in the last couple of decades of studies following current distributional methods such as Behavioral Profiles (cf. Gries and Divjak 2009; Gries 2010) and word-space models (cf. Heylen et al. 2008; Levshina and Heylen 2014), which focus precisely on the semantic structure of specific synonyms. Even though these studies have all successfully uncovered subtle distinctions in meaning between synonyms by analyzing, for instance, their semantic, syntactic, and/or stylistic behavior, they have only dealt with the semantic structure of synonyms from a synchronic viewpoint, while their diachronic evolution generally has been neglected. The present paper seeks to partially

fill this gap by focusing on the distributional patterns of a set of adjectival near-synonyms in the history of American English.

By adopting a distributional orientation, this study examines the attributive uses of four adjectival near-synonyms, namely, *fragrant*, *perfumed*, *scented*, and *sweet-smelling*, in the latter part of Late Modern (LModE) and Present-day American English (PDE) (1850–2009).[2] This is done by drawing on data from the *Corpus of Historical American English* (*COHA*) (Davies 2010-) and by carrying out a Hierarchical Configural Frequency Analysis (HCFA).[3] The focus is on the collocational behavior of the synonyms, and more specifically on the semantic categories of their head nouns.

There are multiple reasons for deciding on this group of synonyms. First, to the best of my knowledge, no previous study on synonymy has focused on lexical items within the domain of smell. In fact, a common claim in the literature is that people are bad at naming smells (cf. Henning 1916; Lorig 1999; or Yeshurun and Sobel 2010; among others), which is probably related to the fact that the odor vocabulary has not been as widely researched as other domains such as that of color (cf. for instance Berlin and Kay 1969; or Biggam 2012) or cookery (e.g. Lehrer 1969, 1974), which have been argued to be much more extensive and precise in English. However, the *Historical Thesaurus of the Oxford English Dictionary* (*HTOED*) lists 894 words under the semantic category SMELL AND ODOR, which have been and/or are used in English, a fact which appears to contradict previous claims. Of these words a substantial amount are adjectives similar to those selected in the present study with some minor variations in definition or use, making this a particular interesting semantic field for the study of near-synonymy. Second, the four synonyms appear to be interchangeable as they are defined in practically the same way, as "having a pleasant smell" in existing reference material (*Cambridge Dictionary*, *Collins Dictionary*, *English Oxford Dictionaries*, *Longman Dictionary of Contemporary*

2 Over the course of this paper, *sweet-smelling* refers to the adjective, whereas SWEET-SMELLING is used to refer to the whole concept, as denoted by the four synonymous adjectives (i.e. *fragrant*, *perfumed*, *scented*, and *sweet-smelling*). Additionally, SMALL CAPITALS are used for semantic categories and conceptual metaphors, while variables and meanings are placed within single quotation marks (‘’). Dictionary citations are structured as follows: (*OED* s.v. *breath* noun 3a). The first element indicates the name of the dictionary in abbreviated form, followed by s.v, which stands for *sub voce* "under the word", head word, part of speech, and finally sense number.

3 The HCFA is a multivariate technique used for categorical variables which analyzes the simultaneous effect of various factors on a particular phenomenon. This test aims at identifying interactions between the variable and variable-levels included in the analysis in order to detect types or groups with similar behaviors or characteristics (Hilpert 2013: 56–57). For more details, see Section 3.2.

English, *Merriam-Webster*, *Oxford English Dictionary* (*OED*), *The American Heritage Dictionary of the English Language*).[4] Additionally, some dictionaries employ circular definitions, i.e. definitions in which the synonyms are defined in terms of each other. Third, the examples of usage provided do not clarify which sort of nouns they commonly co-occur with. Several nouns such as *flower, soap*, and *smell* appear modified by more than one of the adjectives without any apparent change in meaning. Only some of the dictionaries consulted offer information about their particular nuances of meaning: three of the eight dictionaries consulted provide additional senses for *perfumed* (*English Oxford Dictionaries*, *Collins Dictionary*, *OED*) and *scented* (*Collins Dictionary*, *OED*); none for *fragrant* or *sweet-smelling*. According to these dictionaries, there seems to be a difference in nuance in the case of the former two adjectives depending on whether the source of the smell is natural or artificial. This difference is reflected in the kind of nouns that the adjectives modify. In the case of natural sources of smell, *perfumed* and *scented* modify nouns, as for instance, *flower, coffee*, and *sea*, which can release a natural sweet pleasant smell, whereas in the case of artificial sources, they collocate with nouns such as *dress, cream*, and *soap*, which can acquire an agreeable smell only by being impregnated by a sweet-smelling substance. Yet, the information is limited, and it is impossible to identify in which contexts the synonyms can be used interchangeably. This lack of adequate and comprehensive coverage of the semantic properties of the four synonyms calls for a more detailed analysis.

The remainder of this article is organized as follows: Section 2 provides an overview of the theoretical foundations and developments of the distributional method to lexical semantics, and more specifically to the analysis of synonyms. Section 3 focuses on the data retrieval, methodology, and statistical analysis proper, whereas Section 4 puts forward the results obtained as well as a discussion of their implications. Finally, Section 5 offers some concluding remarks and suggestions for future research.

4 The selection of the four near-synonymous adjectives is based on careful examination of the definitions and examples of usage provided by these reference materials. While it is true that some other adjectives share the same definition as the selected ones (e.g. *fragranced*) or are listed as their synonyms (e.g. *aromatic*) and could thus have been chosen as well, these lexical items have been excluded for different reasons. The adjective *fragranced* is classified as rare in several of the reference materials and it is not attested in *COHA*. The definition of *aromatic* differs slightly from the adjectives under scrutiny and does not necessarily imply the trait sweetness. Additionally, many of the other adjectives listed in the HTOED are much less frequent than the four items examined here. Nevertheless, a selection of a wider range of near-synonyms, including adjectives such as *sweet-scented, sweet* (*OED* s.v. n.6), *balmy* (*OED* s.v. n.3), or *redolent*, might prove useful or even necessary in future studies.

2 Literature review

In the last decades, there has been a particular interest in the semantic structure of synonyms, which has primarily been due to scholars who follow a contextual approach to semantics. However, the distributional method for lexical semantics has a long history, going back to the neo-structuralist school. According to this method, the semantic and functional features of lexical items closely correlate with their distributional patterns. More specifically, this tradition rests on three major tenets: (1) its radical usage-based orientation, (2) the paramount importance of collocation, i.e. the "lexical relation between two or more words which have a tendency to co-occur within a few words of each other in running text" (Stubbs 2001: 24), and (3) its strong link with statistical analysis.

Scholars such as John Firth (1957a, 1957b) and John Sinclair et al. (1966, 1987, 2004) were among the first to suggest that distributional patterns and collocations are essential for the analysis of lexical semantics. However, it was not until the latter half of the twentieth century that major advancements in corpus research took place which allowed these scholars' view of thinking to progress considerably. Association measures such as the Pointwise Mutual Information index (PMI) (Church and Hanks 1990) became groundbreaking as they permitted to quantify the probability of association between collocates, and have since been widely employed. Additionally, further statistical techniques such as regression analysis and clustering methods, e.g. hierarchical agglomerative cluster analysis (cf. Gries 2009: 306–319) and conditional inference trees (cf. Levshina 2015: 291–300) have led to further development in the field.

Concerning synonymy, in the 1990s and early 2000s several studies paid close attention to the collocational and colligational patterns of near-synonyms (i.a. Kennedy 1991; Church et al. 1994; Biber et al. 1998: Section 2.6; Partington 1998: Section 2 and 3.6; Kjellmer 1995; Taylor 2003), employing association measures like the PMI or the *t*-score (Gries 2001, 2003).[5] Despite their valuable findings, which have contributed to a better understanding of the behavior of synonyms, there is still room for polishing this method. In fact, Gries (2010) highlights three areas which could be fine-tuned. First, most of the aforementioned studies focus on pairs of synonyms (cf. Kjellmer 1995; on *almost* and *nearly*; Taylor 2003; on *high* and *tall*), whereas the majority of near-synonyms come in larger sets. Second, the range of distributional patterns considered is

5 Colligations are defined as "the collocation of a node word with a particular grammatical class of words" (McEnery et al. 2006: 82).

rather limited, being restricted either to collocational or colligational behavior. Last, some studies lack explanations of the implications of the results obtained and do not relate them to any linguistic or semantic theory.

In view of this backdrop, scholars within the context of cognitive semantics, have recently employed innovative methods to the study of synonymy (see Section 1) with exceptionally positive results, thus proving these methods to be particularly effective in delineating the fine-grained aspects of meaning and in unfolding the differences existing between synonyms. Most research in this domain have concentrated on sets of near-synonymous verbs (cf. Divjak 2006; Divjak and Gries 2006; Speelman and Geeraerts 2009; Levshina 2011). In a Behavioral Profile study, Divjak and Gries (2006) examined the distributional patterns of nine Russian verbs of trying to analyze their syntactic and semantic patterning. The syntactic variables include, for instance, verb related characteristics such as aspect and mode as well as clause and subject related characteristics such as sentence type and clause structure. Semantic variables comprise aspects concerning the nominative subject paradigm, that is, oppositions such as animate vs. non-animate or concrete vs. abstract. Moreover, they analyze the adjacent verb collocates of the synonyms by grouping them into semantic categories such as PHYSICAL ACTION, PHYSICAL PERCEPTION, SPEECH, and INTELLECTUAL ACTIVITY. This was done with the intention of unfolding their primary and secondary meanings and differentiating them from their respective synonyms. Divjak and Gries' study is just one of the studies which have proven the usefulness of the Behavioral Profile approach in facilitating the identification of the semantic structure of verbs and can be said to have paved the way for further and more detailed research along the same lines on other parts-of-speech: adjectives (cf. Gries and Otani 2010; Liu 2010), nouns (cf. Liu 2013), and adverbs (cf. Liu and Espino 2012).

Concerning research on near-synonymous adjectives, we find Gries and Otani (2010) on *big*, *great*, and *large* and their respective antonyms (i.e. *little*, *small*, and *tiny*), and Liu (2010) on *chief*, *main*, *major*, *primary*, and *principal*. Liu delineated the semantic structure of the five synonyms by drawing on data from the *Corpus of Contemporary American English* (Davies 2008–) and carrying out a HCFA. He discovered significant differences between the attributive adjectives by measuring their collocational (i.e. the semantic categories of nouns modified by the synonyms), colligational, and stylistic preferences (i.e. text type). The reasons for paying close attention to the head nouns are well justified, since the existing literature has demonstrated that looking at the nouns with which adjectives collocate is one of the best ways to reveal the nature of the semantics of adjectives. For instance, Geeraerts' (1986) study on the Dutch adjective *vers* "fresh" demonstrated that the fine-grained aspects of meaning of polysemous

adjectives can be discovered by examining their head nouns. Geeraerts provides evidence for *vers* having different meanings depending on the nouns it accompanies. For instance, when it modifies *wound*, *vers* means "recent", but when it modifies the noun *air*, it takes on a slightly different meaning, namely "pure, untainted", or "optimal". Gries' (2001, 2003) studies on English *–ic* and *–ical* adjective pairs, e.g. *economic* and *economical*, has also contributed interesting findings by focusing on the head nouns of the paired adjectives. Liu (2010), who classified the head nouns into six semantic categories, to wit CONCRETE, ABSTRACT, DUAL, INSTITUTION, POSITION TITLE, and NON-POSITION TITLE, learnt that whereas dual and abstract nouns were often modified by the five adjectives, the other semantic categories tended to be dominated by one or two of them: for instance, while concrete nouns preferred *main*, position title nouns were dominated by *chief*.

As demonstrated above, multiple studies have been carried out with effective results, particularly within the context of cognitive semantics. In fact, this theory converges with the original neo-structuralist distributional method in many ways, not the least due to their reliance on actual usage-based data and their quantitative orientation. Given the rapid increase of studies in cognitive semantics over the last two decades, its importance and utility cannot be underestimated (cf. Geeraerts 2010: 264–265).

This brief review has demonstrated the effectiveness of the distributional corpus-based approach in uncovering fine-grained semantic and usage distinctions between sets of synonyms as well as the motivations for continuing with research along the same lines. This is precisely what the present study aims at, and the findings obtained offer a significant contribution to the field. It does so by considering the semantic category of head nouns, and by presenting a more detailed semantic classification than most previous studies. More importantly, the approach adopted differs from earlier research in that the diachronic development of the synonyms is also accounted for, something which to my knowledge has not been done before. The main research questions addressed are:

(1) Do the near-synonyms appear with diverse semantic categories of nouns? If so, are the differences significant?
(2) Are there any major changes taking place over the time span analyzed which have an influence on the use of the concept SWEET-SMELLING? If so, where are these changes taking place and what are the implications?
(3) Could the innovative methods put forward in the context of cognitive semantics be applied to the diachronic study of near-synonyms with equally effective results as in previous synchronic research?

3 Data and method

To provide a detailed description of the synonyms selected, *fragrant, perfumed, scented*, and s*weet-smelling*, the first and most elementary information required is their overall frequency. However, this study focuses only on the attributive uses, and more specifically on those uses in which there is no intervening material between the synonymous adjectives and the head nouns. Therefore, it was necessary to discard the uses where the synonyms function as verbs or as predicative complements, as well as the attributive uses in which the head nouns did not directly follow the adjectives.[6] Consequently, this study is based on 1,666 instances of the near-synonyms. This section describes the process of the data retrieval, the semantic classification of nouns (Section 3.1), and the statistical analysis (Section 3.2).

3.1 Corpus: data retrieval and semantic classification

The data used was retrieved from *COHA*, which, with approximately 400 million words, suits perfectly the purposes of this analysis, as it allows us to examine low frequency items such as the synonyms under study with more reliable results than smaller corpora. However, even though *COHA* covers the time span 1810–2009, this study only focuses on the decades from 1850 onwards. The reason for this is that *COHA* is not a balanced corpus in terms of number of words, as it contains considerably more data for the later decades than for the earlier ones. In fact, until the 1850s there are less than 16,000,000 words per decade, whereas the last two decades, 1990s and 2000s, include nearly double the amount of running text. To make the decades comparable, this analysis is based on approximately 16,493,826 words from each decade (i.e. the same as the total number of the first analyzed decade, the 1850s), resulting in a total number of 263,901,889. Finally, the decades were grouped into four 40-year periods, that is, 1850–1889 (Period 1: P1), 1890–1929 (Period 2: P2), 1930–1969 (Period 3: P3), 1970–2009 (Period 4: P4), of around 65,975,466 words each, as shown in Table 1.

---

6 Examples as the following were thus not considered in the analysis:

(i) *perfumed* as verb: *He held the shirt up. It* ***perfumed*** *the air rapidly*. (COHA, 1943, FIC)

(ii) predicative use: *They were* ***sweet-smelling****, sweet as flowers out of season.* (COHA, 2001, FIC)

intervening material: *I had picked my mother some* ***fragrant*** *pale pink Renaissance roses from her garden.* (COHA, 2009, FIC)

Table 1: Number of words per decade and period.

P 1	Words	P 2	Words	P 3	Words	P 4	Words
1850	16,493,826	1890	16,493,638	1930	16,493,849	1970	16,493,820
1860	16,493,824	1900	16,494,074	1940	16,493,903	1980	16,493,824
1870	16,493,817	1910	16,493,958	1950	16,493,875	1990	16,493,914
1880	16,493,999	1920	16,493,776	1960	16,493,893	2000	16,493,899
Total	**65,975,466**	**Total**	**65,975,446**	**Total**	**65,975,520**	**Total**	**65,975,457**

Total words in subcorpus: 263,901,889

To acquire information concerning the nuances of meaning of the synonyms, it was necessary to identify the specific usage patterns of each adjective. To do so, I followed the methodology employed in Liu (2010). As the focus lies on the head nouns, the most common noun collocates of each adjective were retrieved by using the PMI score, which compares the likelihood of two words co-occurring with the likelihood of the individual words (Church and Hanks 1990: 23).[7] This procedure facilitated the creation of a list of approximately ten top adjacent nouns, that is, R1 collocates, for each of the synonyms in the four different periods established.[8] Given that I queried the PMI in each of the periods established, I ended up with a total of 16 lists (i.e. four for each synonym, one for each period). However, as there was a considerable degree of overlap between the lists, both between the adjectives and the different periods, the final list of top nouns contained 59 different noun types in total. These 59 nouns were subsequently grouped into nine fine-grained semantic categories based on the semantic classification offered by the *HTOED*, which classifies each of the senses of a given word into categories: FOOD & DRINK, PLANTS & FLOWERS, EARTH, BODY, MATTER, SENSATION, AESTHETICS, CLEANING, and TEXTILE & CLOTHING.[9] Then, a query was conducted in *COHA* to retrieve all adjacent

7 The PMI score was used since there is a specific query in *COHA* which calculates the PMI of the collocates of the target word. The PMI was used is it favors words with a generally low frequency in the corpus, but which tend to collocate with the adjectives. Additionally, The PMI was used instead of the more habitual *t*-score, since the PMI is more suitable when dealing with content words as is the case in the present analysis, while the *t*-score performs better in the case of function words (cf. Church et al. 1994 for further details on these association measures).

8 In view of the low frequency of some of the adjectives, especially *sweet-smelling*, it was not always possible to obtain a list of 10 top nouns for each period. For this reason, I also made used of the frequency measure. This measure provides the nouns that the adjectives most frequently modified irrespective of their overall frequency in the corpus. As Liu (2010: 63) points out, "the frequency measure and the MI score can complement each other in better identifying the nouns that are typically modified by the adjectives."

9 As each of the synonyms modify a huge number of individual types of noun, it would be extremely time consuming to categorize all senses of the individual head nouns. Consequently,

head nouns of the synonyms. In this case, all nouns directly to the right (R1) of the node word (i.e. the synonyms) were retrieved, but only those that could be classified into any of the nine semantic categories were selected and included in the analysis, thus increasing the total number of head nouns examined to 110. Table 2 lists these categories together with the 110 nouns belonging to each of them.

Table 2: Classification of the nouns (lemmas) modified by the synonyms.

Category	Nouns
1. FOOD & DRINK	*apple, beverage, bread, chicken, coffee, cup, drink, food, fruit, liquid, meal, omelet, rice, spice, tea, wine*
2. PLANTS & FLOWERS	*bloom, blossom, bower, flower, garden, geranium, grass, herb, lip*, leaf, petal, pine, rose, shrub, vine, violet*
3. EARTH	*breeze, brook, dew, flood, gale, grove, hill, sea, vale, valley, wind*
4. BODY	*arm, breath*, cheek, face, flesh, hair, hand, head, limb, lip*, lock, mouth, shoulder, skin, wrist*
5. MATTER	*air, atmosphere, candle, cloud*, dust, fume, gas, oil*, night, smoke, steam, vapor*
6. SENSATION	*aroma, breath*, flavor, incense, scent, smell, odor, taste*
7. AESTHETICS	*cologne, cosmetics, cream, oil*, ointment, powder, talcum, wax*
8. CLEANING	*deodorant, dish-water, disinfectant, napkin, soap, soap-powder, sponge, spray, suds, tissue, wash-ball*
9. TEXTILE & CLOTHING	*blanket, cambric, cloth, dress, flannel, garment, glove, lace, linen, pillow, robe, sheet, shirt, silk*

As shown in Table 2, some of the nouns were classified into more than one category (i.e. those nouns followed by an asterisk, for instance, *breath*, *lip*, and *oil*), since the adjectives functioned as modifiers of several senses of the same noun. The following examples of *perfumed* and *fragrant* serve to illustrate this point:

(1) *Close by her side I rested, and once when she stooped toward me in the telling of the tale, I felt her* ***perfumed breath*** *upon my cheek.* (COHA, 1934, FIC)

(2) *And they passed under a shadowed grape-arbour and past a dead locust-tree, which a vine had made into a green tower of waving tendrils, and from which came the* ***fragrant breath*** *of wild grape [...]* (COHA, 1900, FIC)

While *breath* in (1) refers to "the air exhaled from the lungs" (*OED* s.v. *breath*, noun 3a), in (2) it refers to "the air exhaled from anything, or impregnated with

the list of top nouns was created solely to facilitate the identification of the categories of nouns that the adjectives most frequently modify.

its exhalations, and retaining its characteristic odour" (*OED* s.v. *breath*, noun 2b). These two senses of *breath* were classified into two distinct categories in the *HTOED*, namely BODY (sense 3a) and SENSATION (sense 2b). Moreover, some classifications need to be further clarified. This is the case of metonymic and metaphoric uses, as in examples (3) and (4), as well as nouns like *garden*, illustrated in (5), which I thought to be misclassified in the *HTOED*.

(3) *They enjoy afternoon tea precisely as other women enjoy it, and over a* ***fragrant cup*** *they relax, talk and laugh and tell stories* [...] (COHA, 1905, NF)

(4) *Down in the vale where cowslips are growing, Where violets breathe thro' sweet* ***scented lips****, Where brook o'er the bright pebbly bottom is flowing*, [...] (COHA, 1895, FIC)

(5) *And the moon comes through the tree-tops in splashes, and there is softness and a shade, and it is all like a* ***scented garden*** *in some old Arabian story, and the senses are affected and, maybe, the reason*. (COHA, 1892, FIC)

In (3), *cup* is an example of an extension of meaning through the conceptual metonymy CONTAINER FOR CONTENTS (cf. Kay and Allan 2015: 165). Thus, the meaning of *fragran*t does not apply to the cup itself, but to the content within. (4) is a typical example of metaphorical sense extension, in which *lip* is used to refer to a part of a flower. The lemma *garden* is classified into FOOD & DRINK in the *HTOED*, but I decided to include it in PLANTS & FLOWERS instead, in view of examples like (5), in which *scented* refers to the pleasant smell characteristic of the plants and flowers present in the garden.

3.2 Statistical analysis: HCFA

To establish whether the independent variables have a significant effect on synonym choice, the data was fed into a HCFA analysis by applying Gries' (2004) *HCFA 3.2* script for *R* (R Core Development Team, 2015). The database contained, for each relevant observation of the four synonyms extracted from *COHA*, a tag signaling the semantic category of the particular head noun following the synonym as well as a tag indicating the period to which the instance belongs. The HCFA is a multifactorial test which "explores every possible combination of variable levels [...] for the presence or absence of an effect" (Gries 2009: 240). It is an extension of the chi-square test and allows for the inclusion of more than two variables, thus suiting the present analysis, which considers the effect of the variables "Period" and "Semantic category of head noun" on "Synonym". The test calculates which variables and variable-

levels interact in significant ways by comparing their observed and expected frequencies. The results show which cell frequencies in the contingency table are significantly higher or lower than what is expected by chance by grouping them into *types* (T) and *antitypes* (A); those cells which are significantly higher are types and those which are significantly lower are antitypes (Hilpert 2013: 55–66).[10] The output includes the global significance value of the table together with the configurations that account for it. To illustrate the point, Table 3 shows an extract of the results offered by the HCFA test.

Table 3: Extract of the results of the HCFA.

Synonym	SEM.CAT	Period	Freq.	Exp.	Cont.chisq	Obs-exp.	Dec	Q
Fragrant	P&F	1850–1889	228	122.501	90.8567	>	***	0.068
Fragrant	P&F	1890–1929	127	102.4222	5.8978	>	ns	0.016
Fragrant	P&F	1970–2009	33	59.4252	11.7508	<	*	0.016
Fragrant	Cl	1850–1889	1	23.5267	21.5692	<	***	0.014
Perfumed	P&F	1890–1929	10	32.1993	15.305	<	***	0.014
Scented	Cl	1970–2009	27	4.0777	128.855	>	***	0.014
Scented	Ma	1970–2009	33	10.7217	46.2914	>	***	0.013
Perfumed	P&F	1850–1889	19	38.5116	9.8854	<	*	0.012
Fragrant	Cl	1890–1929	1	19.6705	17.7213	<	***	0.011
Perfumed	P&F	1970–2009	1	18.6819	16.7354	<	***	0.011

The first three columns specify the levels of each variable, i.e. *fragrant*, *perfumed*, *scented*, and *sweet-smelling* for "Synonym"; FOOD & FRINK (F&D), PLANTS & FLOWERS, (P&F), BODY (BO), EARTH (EA), MATTER (MA), TEXTILE & CLOTHING (T&C), SENSATION (SE), AESTHETICS' (AE), and CLEANING (CL) for "Semantic Category of head noun" (Sem.Cat); and 1850–1889 (P1), 1890–1929 (P2), 1930–1969 (P3), and 1970–2009 (P4) for "Period". The next two columns, "Freq" (Frequency) and "Exp" (Expected), give information about the observed (Freq) and expected (Exp) frequencies of each configuration in the table, while the "Cont.chisq" (Contribution to chi-square) column provides their respective chi-square values. "Obs-exp" (Observed-expected) displays the relation between "Freq" and "Exp", "<" meaning less observed frequency than expected, and ">" more than expected. The column "Dec" (Decision) reveals whether the configurations are statistically significant or not: "ns" = not significant, "ms" = marginally significant, "*" = significant at the 0.05 level,

10 In Tables 4, 5, 6, 7, and Table 8, which present the results of the analysis, those cells in which the figures are followed by a T are types, whereas those followed by an A are antitypes. Cells which do not contribute significantly are followed neither by a T nor an A.

"**" = significant at the 0.01 level, "***" = significant at the 0.001 level. Last, "Q" (Coefficient of Pronouncedness) indicates the size of the effect; the higher, the stronger. Additionally, HCFA also gives the global chi-square and significance values for the whole table, which for this example are the following, and thus provide evidence of its significance: $\chi^2 = 1003.043$ (d.f. = 129) and $p < 0.001$.

4 Results and discussion

This section presents the results obtained and a discussion of their methodological and theoretical implications. Section 4.1 deals with the overall frequency of the concept SWEET-SMELLING as well as that of the individual synonyms, but also with the changes in frequency over the time span analyzed. Section 4.2 provides the general collocational behavior of the four adjectives, whereas Section 4.3 gives an account of the changes in collocational patterns, both of the concept and of those specific to each synonym. Finally, in Section 4.4 a general discussion of the findings is offered.

4.1 Overall frequency and individual frequency development

A total number of 1,666 tokens of the four synonymous adjectives were analyzed, distributed as shown in Figure 1. Starting with the overall frequency, the findings point to a substantial difference among the adjectives: *fragrant* is much more frequent than the rest, *perfumed* and *scented* are quite similar, competing for the second position, and *sweet-smelling* is the least frequent. In fact, the frequency of *sweet-smelling* is only about 6.5%, that of *perfumed* 17.6%, that of *scented* 20%, and that of *fragrant* 55.9%. However, this distribution, as will be shown, varies greatly over the different periods.

Concerning the changes in the frequency of the near-synonyms, *fragrant* and *perfumed* show a general downward tendency, which is especially pronounced in the case of *fragrant*, whereas *scented* and *sweet-smelling* increase over the periods, especially in the last two. The statistical analysis shows that some of the findings are significant as seen in Table 4: while *fragrant* goes from being a type in P1 to an antitype in P4, *scented* and *sweet-smelling* go from being antitypes in P1, to types in P3 and P4, respectively. In fact, toward the last two periods, the adjectives become more similar in terms of frequency, thus displaying a gradual convergence over the time-span analyzed (see Figure 1). To illustrate this fact, in P1, the distribution of the adjectives expressed in percentages is as follows: *fragrant* accounts for 67.4% of the use of concept SWEET-SMELLING, *perfumed* for

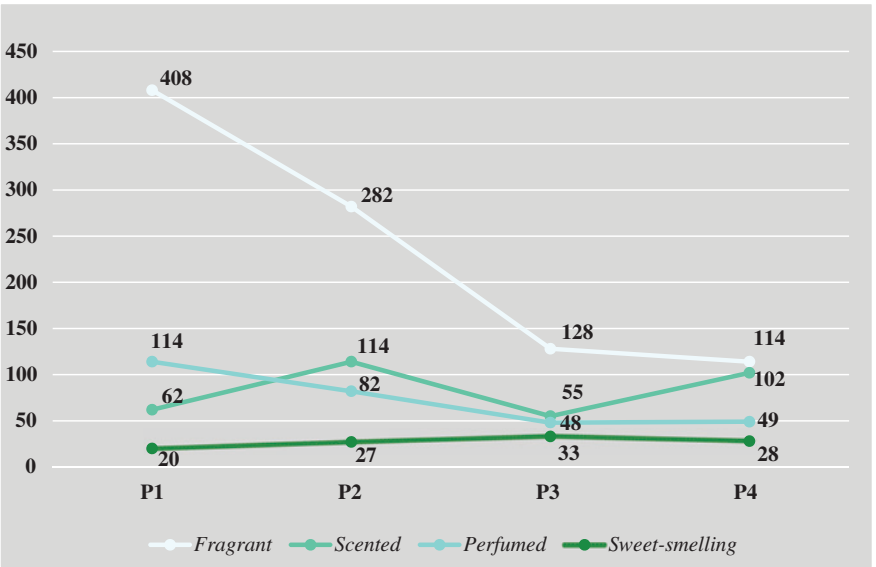

Figure 1: Frequency of the adjectives throughout 1850–2009.

Table 4: Frequency of the adjectives throughout 1850–2009.

	Adj.				
Period	***Fragrant***	***Scented***	***Perfumed***	***Sweet-smelling***	**Total**
P1: 1850–1889	408 T	62 A	114	20 A	604
P2: 1890–1929	282	114	82	27	505
P3: 1930–1969	128	55	48	33 T	264
P4: 1970–2009	114 A	102 T	49	28	293
Total:	932	333	293	108	1,666

18.9%, *scented* for 10.3%, and *sweet-smelling* for 3.3%. In contrast, the figures for P4 show a totally different picture, with *fragrant* making up solely 38.9% of the uses (i.e. more than 30% less than in P1), *perfumed* 16.7%, whereas *scented* stands for a whole 34.8% (that is, nearly 25% more than in P1) and *sweet-smelling* for 9.6%, i.e. 6.3% more than in P1.

These data demonstrate that the distribution of the near-synonyms, representing SWEET-SMELLING, is changing over the time-span analyzed. First, considering the total number of occurrences of the concept, there is a global

downward tendency, from 604 in P1 to 293 in P4 (see Table 2). Second, whereas some of the adjectives decrease considerably in frequency, other increase significantly. Third, there exists a merging in frequency in the two later periods. This indicates that the concept is not used as much in PDE as in the latter part of LModE. Due to this general decrease it is not surprising that the lexical items denoting the concept also decrease in use. However, the fact that only two of the adjectives (i.e. *fragrant* and *perfumed*) do so, whereas the other two (i.e. *scented* and *sweet-smelling*) increase, calls for further explanation. Analyzing the adjectives in different contexts and with diverse categories of nouns might shed some valuable light on this evolution.

4.2 General collocational usage patterns throughout 1850–2009

As seen in Figure 2, although there exist significant differences between the synonyms, there are also some similarities. Figure 2 plots the proportion of cases of each synonym in relation to the others in each of the nine levels of the predictor "Semantic category". On the vertical axis we have the four synonyms, each one represented by a different color, while the semantic categories are plotted on the horizontal axis. The size of each of the partitions in the figure shows the proportion of cases of a particular configuration, in comparison to the rest: whereas the width of the columns reflects the amount of cases included in each semantic category, the height of the divisions on the vertical axis represents the proportion of instances of the synonyms within each semantic category. To illustrate this fact, the top left box in Figure 2 represents the proportion of FOOD & DRINK nouns modified by *fragrant*.

First, it can be appreciated that some of the noun categories occur much more frequently (e.g. MATTER and PLANTS & FLOWERS) with the attributive adjectives than others (e.g. FOOD & DRINK, EARTH, and AESTHETICS). For instance, MATTER nouns (e.g. *air* and *steam*) are rather prominent with all four adjectives, although none of them shows up as a type with these nouns. In fact, this semantic category is either the most or the second most frequent in all cases: with *perfumed* it is the most frequent, accounting for 22.9% of its uses, and with the other three adjectives it is the second most frequent, making up for 15.9% with *fragrant*, 22.5% with *scented*, and 13.9% with *sweet-smelling*. Similarly, in some of the less common categories of nouns, all four or at least three of the adjectives seem to be relatively evenly distributed, and there are no statistically significant differences between them. This is the case of EARTH (e.g. *hill* and *sea*) and SENSATION (e.g. *smell* and *flavor*), except for *scented* in the latter category, which here turns out to be an antitype (see Table 5).

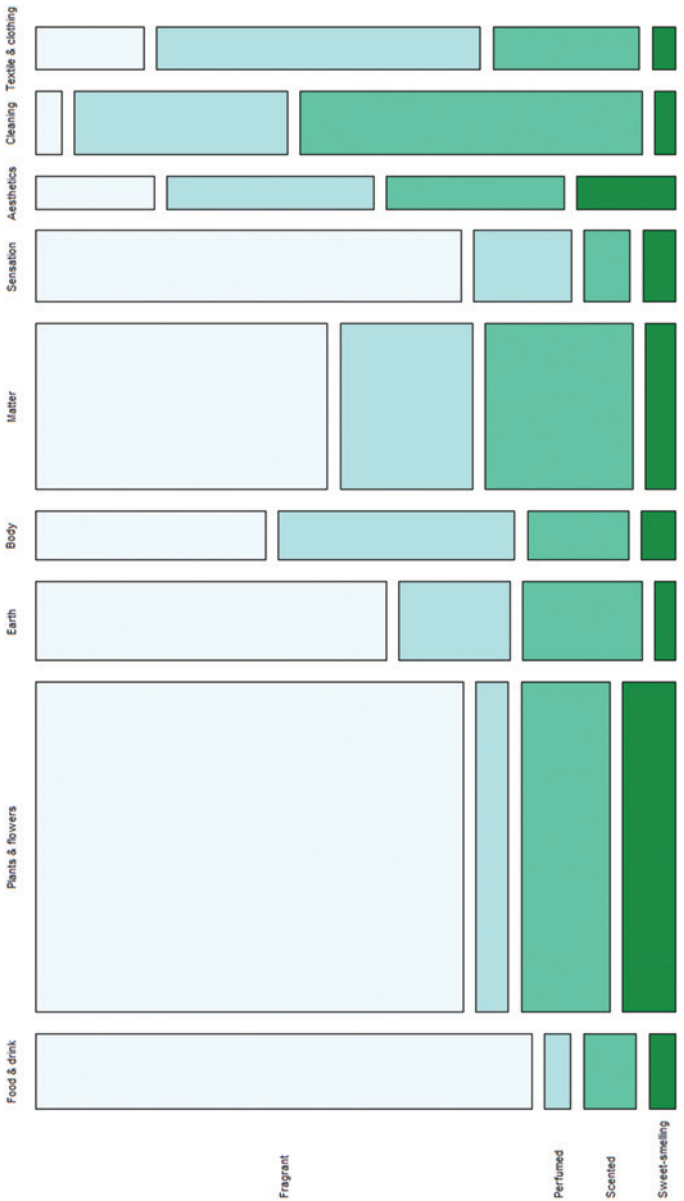

Figure 2: Distribution of the synonyms across the semantic categories.

Table 5: Distribution of the adjectives across the nine semantic categories.

	Adj.			
Cat.	***Fragrant***	***Scented***	***Perfumed***	***Sweet-smelling***
FOOD & DRINK	113 T	12 A	6 A	6
PLANTS & FLOWERS	429 T	89 A	33 A	53
BODY	34	15	35 T	5
EARTH	85	29	27	5
MATTER	148	75	67	15
TEXTILE & CLOTHING	14 A	19	42	3
SENSATION	92	10 A	21	7
AESTHETICS	12 A	18	21 T	10
CLEANING	5 A	66 T	41 T	4
Total:	932	333	293	108

However, the remaining semantic categories are dominated by one or two of the adjectives and the results do reveal significant differences between the synonyms. In FOOD & DRINK (e.g. *chicken* and *wine*), *fragrant* is the only type and there are just a few examples of the other three adjectives with these nouns. What is more, both *scented* and *perfumed* are antitypes with nouns denoting food and drink. Likewise, *fragrant* is the only type with PLANTS & FLOWERS (e.g. *geranium* and *leaf*) whereas *scented* and *perfumed* are both antitypes. The only difference is that *sweet-smelling* is a considerably frequent modifier of this sort of nouns, with nearly half of its examples (49%) occurring with nouns designating plants or flowers. Contrarily, in the CLEANING category (e.g. *disinfectant* and *soap*), *scented* and *perfumed* are the only types while *fragrant* turns out to be an antitype with only five examples (that is, approximately 0.54% of its use). Similarly, *sweet-smelling* is rather infrequent with this kind of nouns as it only occurs collocating with them in four of the instances (3.7%) in the data analyzed. In the remaining categories, namely BODY (e.g. *skin* and *hair*). TEXTILE & CLOTHING (e.g. *dress* and *silk*), and AESTHETICS (e.g. *cream* and *lotion*), *perfumed* prevails, being the only type, while *fragrant* is considered an antitype in the latter two.

The findings obtained show that the adjectives do, in general, have different collocational preferences as they tend to co-occur with different categories of nouns. These differences in collocational affinity demonstrate that synonyms are normally not interchangeable, but rather tend to collocate with diverse types or sets of words despite sharing the same sense (Cruse 1986: 278–281). For instance, Murphy (2003: 256) explains that although some of the existing differences between near-synonyms, for instance those relating to collocations or frequency, do not influence the communicative intent, they do provide evidence that

synonyms are rarely substitutable. Moreover, she adds that words typically collocate with a limited amount of other words and, despite the fact that replacing a word by any of its synonyms does not always lead to ungrammaticality, it normally results in a non-natural combination due to collocational affinities.

One important picture emerges from the results obtained so far, namely a clear division of semantic labor between the synonyms regarding the two senses previously referred to (cf. Section 1): (1) "Naturally possessing or emitting a pleasant sweet smell" and (2) "permeated with a substance which artificially adds a pleasant sweet smell". As seen in Table 2, by closely analyzing the nouns in each semantic category, it becomes clear that some of them solely comprise nouns which denote entities with a natural pleasant smell, i.e. PLANTS & FLOWERS, FOOD & DRINK, and EARTH, whereas other include nouns which exclusively refer to man-made or artificial objects, i.e. TEXTILE & CLOTHING, AESTHETICS, and CLEANING. Whereas *fragrant* is a type with nouns which denote entities with a natural pleasant smell, *perfumed* and *scented* are types with nouns referring to man-made, artificial objects, i.e. those in CLEANING (in the case of both adjectives), AESTHETICS, and TEXTILE & CLOTHING (only in the case of *perfumed*). This division of semantic labor hence points to the fact that *perfumed* and *scented* might be the only adjectives which are truly polysemous, being used also in sense 2, as pointed out by some of the dictionaries consulted (cf. Section 1).

The rest of the semantic categories, i.e. MATTER, BODY, and SENSATION, encompass nouns which can be used with the adjectives either in sense 1 or 2 depending on the context of use. There are certainly instances of the adjectives with these nouns in which it is clear whether they are used in sense 1 or 2, as illustrated in (6) to (8).

(6) *Shadow lAke's citrus-based laundry detergent, Citra-Suds, makes short work of both those problems ($7.99). The light fruity smell lends laundry a* ***fragrant aroma*** *without being perfumey, and its cleansing power is dramatically boosted with an oxygenating agent called Oxi – Clean.* (COHA, 2001, MAG)

(7) *His leaf-land is fair and so marvellous fair, His palm-land is filled with a* ***perfumed air*** *Of magnolia blooms to its dome of blue. His rose-land has arbors of moss-swept oak, – Gray, Druid old oaks; and the moss that sways And swings in the wind is the battle-smoke Of duellists, dead in her storied days.* (COHA, 1887, FIC)

(8) *Things she told me, sitting on the edge of my bed at night: "I was born with a caul. That means I have a sixth sense." Or, touching her* ***perfumed wrist*** *to my cheek: "This is called Shalimar."* (COHA, 2006, FIC)

In these examples, the surrounding context, either intra– or extralinguistic, indicate in which sense the adjectives are used. In (6), the words in the surrounding context window of *fragrant* (e.g. *laundry*) indicate that this instance of the adjective with the sensation noun *aroma* is used in sense 2. In (7), the prepositional phrase *of magnolia blooms* and other collocates, such as *leaf-land* and *arbors*, reveal that *air* denotes the atmospheric air of outdoors (as opposed to the air of confined spaces) and is thus naturally pleasant smelling. Likewise, in (8), *perfumed* is used in sense 2, as made evident by the extralinguistic situation, since one of the characters refers to a well-known perfume, namely *Shalimar*, and general knowledge: people typically put perfume on the wrists.

However, some occurrences of these categories of nouns are more ambiguous and it is difficult – sometimes even impossible – to know for sure whether the modifying adjectives are used in sense 1 or 2. Consider the following two examples, in which *scented* and *fragrant* modify a noun belonging to MATTER and BODY, respectively.

(9) *I climbed back inside and took out the magazine and studied Miss July by flashlight and thought about how her skin would feel where her breasts curved back into her body toward her underarms, and how her nipples would feel, and the hair between her legs, the lips between her legs. When I had finished my business, I lay on top of the sheets and let the warm* ***scented air*** *wash over me.* (COHA, 1995, FIC)

(10) *Off came her earrings, then her clothes, without embarrassment. Breasts or none, her woman's wealth satisfied him. When he touched her nipples the effect was electric. She could be a little wild in bed. "Your fingers are fires." He enjoyed her long body, the cool flesh, sex smell,* ***fragrant hair****.* (COHA, 1961, FIC)

In (9), it is clear that *air* refers to the air of a confined space, but it is not clear whether *scented* denotes a natural odor as emitted by, for instance, the body of the person in the room, or by an artificial one as emitted from, for example, detergent that the sheets mentioned have been impregnated with, or from perfumes worn by him/her. Similarly, in (10), the man might find the hair of the woman naturally fragrant, or its sweetness might be due to the shampoo or another cleansing agent. In both examples, it might even be the case that the pleasant odor results from a mixture of both naturally and artificially pleasant-smelling substances. Occurrences of the adjectives such as these cannot be categorized clearly into either sense 1 or 2 but are open to interpretation. Thus, in these three categories, there are examples in which the adjectives cannot be said to be used clearly in either sense 1 or 2 as in the previous cases (i.e. FOOD & DRINK, PLANTS & FLOWERS, EARTH, CLEANING, AESTHETICS, and TEXTILE & CLOTHING).

The results presented in this section reveal important insights about the phraseological tendencies of the synonymous adjectives, as well as the similarities and differences between them. Notwithstanding, these results do not account for their distribution across the nine categories of nouns at specific time periods during the 160 years considered, for potential changes in collocational behavior, for those specific to each adjective, or for those concerning the concept of SWEET-SMELLING. This is the concern at issue in the next section, in which the collocational preferences of the near-synonyms are analyzed in four separate 40-year periods: 1850–1889, 1890–1929, 1930–1969, and 1970–2009.

4.3 Semantic development: Changes in collocational behavior

To examine the collocational patterning of the synonymous adjectives in the four periods distinguished, it was necessary to cross-cut the three variables "Synonym", "Semantic category of head noun", and "Period" in the HCFA. Due to the relatively low frequency of the examined adjectives, it was difficult to obtain statistically significant results. However, it has been possible to determine which cells in the contingency table contribute to the significance of the results of the general collocational preferences. For instance, both *perfumed* and *scented* showed up as antitypes with nouns belonging to PLANTS & FLOWERS in the general analysis (see Section 4.2), but the results obtained when including "period" reveal that while *perfumed* is an antitype in all four periods with these nouns, only P3 (1930–1969) plays a part in *scented* being considered an antitype (cf. Table 6).

Table 6: PLANTS & FLOWERS throughout 1850–2009.

	Adj.				
Period	***Fragrant***	***Scented***	***Perfumed***	***Sweet-smelling***	**Total**
P1: 1850–1889	228 T	19 A	29	12	288 T
P2: 1890–1929	127	10 A	35	14	186
P3: 1930–1969	41	3 A	5 A	18 T	67 A
P4: 1970–2009	33 A	1 A	20	9	63 A
Total:	429	33	89	53	604

Additionally, this analysis yields findings which ascertain that some major changes in collocational behavior are going on. First, PLANTS & FLOWERS nouns become substantially less frequent over time with *fragrant*, *perfumed*,

and *scented*. In the case of *fragrant*, this category goes from being modified more frequently than expected in the first two periods (a type in P1), to being modified less frequently in the later periods, and finally becoming an antitype in P4. With both p*erfumed* and *scented*, these nouns always appear as collocates less frequently than expected, but a decline can still be appreciated. With *perfumed*, PLANTS & FLOWERS is an antitype throughout 1850–2009, but its percentage of use with the adjective decreases over time, going from 16.7% in P1 to 0.13% in P4. With *scented*, the category becomes an antitype in P3, but there is an overall downward tendency, going from making up 46.8% of the uses in P1 to 19.6% in P4 (cf. Figure 3). The only exception is *sweet-smelling*, which modifies PLANTS & FLOWERS significantly more frequently than expected in P3, thus being a type.

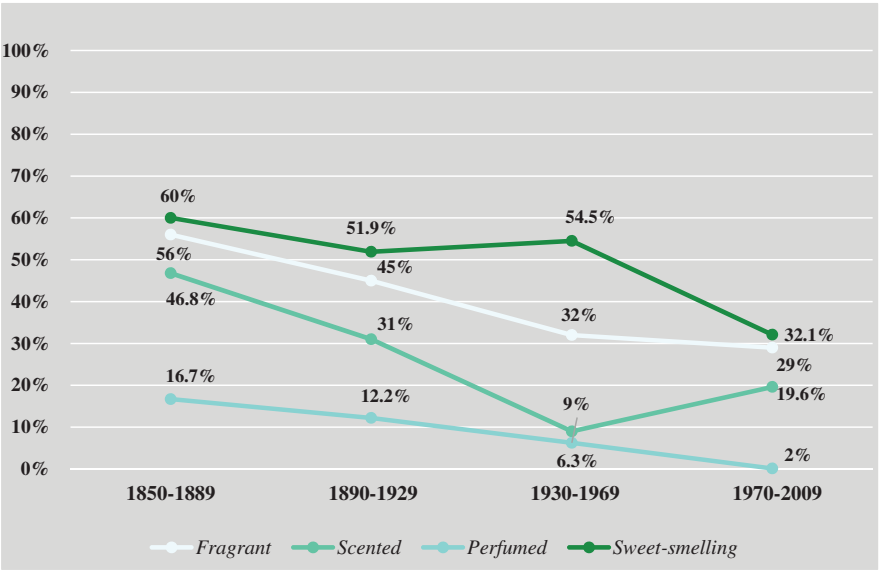

Figure 3: PLANTS & FLOWERS over time.

The output of the HCFA also indicates a change in the categories CLEANING with both *perfumed* and *scented*, and AESTHETICS with *perfumed* (See Tables 7 and 8). The adjective *perfumed* always modifies nouns belonging to CLEANING more frequently than expected, but gains more ground in P3, in which it becomes a type. The percentages of use also indicate a gradual increase as this category goes from totaling 7% of the use of *perfumed* in P1 to 20.4% in P4 (cf. Figure 4). With *scented*, CLEANING nouns occur less frequently than expected in P1, but more frequently in the remaining three periods, becoming a type in P3 and P4.

Table 7: CLEANING over time.

	Adj.				
Period	***Fragrant***	***Perfumed***	***Scented***	***Sweet-smelling***	**Total**
P1: 1850–1889	1 A	8	4	1	14 A
P2: 1890–1929	1 A	12	17	2	32
P3: 1930–1969	1 A	11 T	18 T	1	31
P4: 1970–2009	2 A	10	27 T	0	39 T
Total:	5	41	66	4	166

Table 8: AESTHETICS over time.

	Adj.				
Period	***Fragrant***	***Perfumed***	***Scented***	***Sweet-smelling***	**Total**
P1: 1850–1889	3	3	0	0	6 A
P2: 1890–1929	1 A	3	5	2	10
P3: 1930–1969	3	7	5	6 T	21 T
P4: 1970–2009	5	9 T	8	2	24 T
Total:	12 A	21 T	18	10	61

In fact, these nouns account for only 6.4% of the use of *scented* in P1, whereas it makes up for 26.5%, approximately four times more, in P4 (cf. Figure 4). Something similar goes on with *perfumed* with AESTHETICS nouns: it goes from modifying these nouns less frequently than expected in P1 and P2, to being more common in P3, although not significantly so, to finally becoming a type in P4. As seen in Figure 5, which shows the percentages of use of this adjective with AESTHETICS nouns, there is a rather drastic increase from P3 onwards, going from 2.6% in P1 to 18.4% in P4.

Some important changes are going on as concerns the collocational patterning of the adjectives and, what is more, they are particularly interesting given that the decreases and increases take place in categories containing examples of the adjectives used exclusively in one of the senses: PLANTS & FLOWERS, including sense 1 uses only, and CLEANING and AESTHETICS, including sense 2 uses only. This could point to a decrease of the adjectives in sense 1, and an increase in sense 2. Nonetheless, this interpretation is rather tentative due to lack of more significant results of the individual changes of each of the four synonyms. Therefore, to further investigate this issue, the collocational patterns

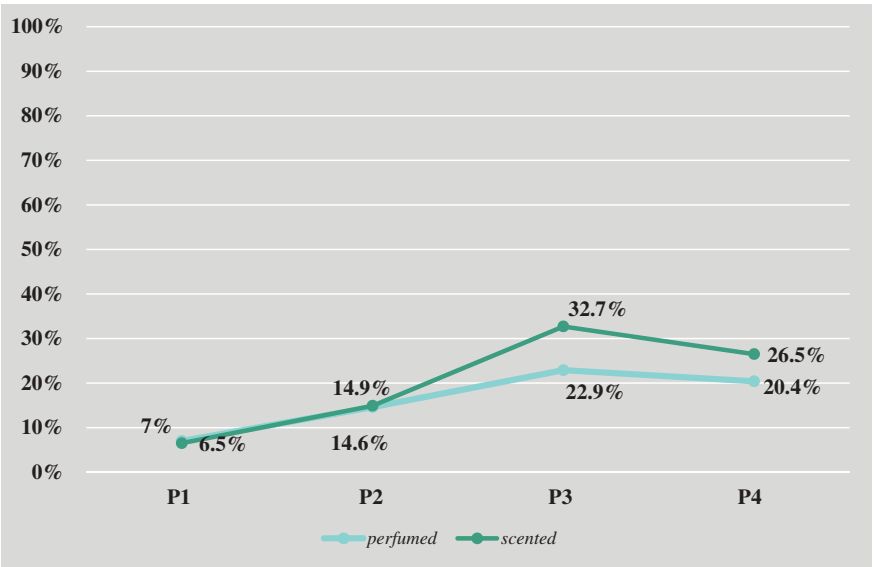

Figure 4: *Perfumed* and *scented* with CLEANING over time.

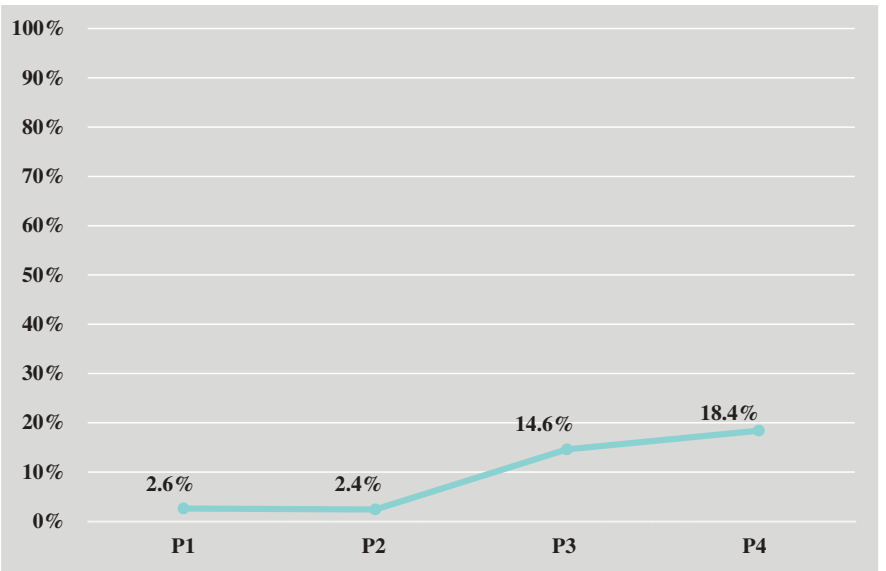

Figure 5: *Perfumed* with AESTHETICS over time.

of the whole concept SWEET-SMELLING, as designated by all four adjectives, were analyzed. A further HCFA test was carried out with the two independent variables "semantic category" and "period", since this configuration permits to identify the collocational patterning of the concept with the different categories across the four periods.

This last set of results reveals the same tendencies for the whole concept SWEET-SMELLING as those obtained for the individual changes of each synonym, namely that PLANTS & FLOWERS goes from being an antitype to becoming a type, and that both CLEANING and AESTHETICS nouns go from being antitypes to becoming types (cf. Tables 6, 7, and 8). Additionally, another significant result appears concerning the category BODY. These nouns go from being modified less frequently than expected by the concept (although not significantly so) in the first two periods to becoming more frequent in the last two periods, being a type in P3. In the remaining semantic categories there are no statistically significant results. However, there are cases in which there exists a change in the relation between observed and expected frequency (see Section 3.2), that is, in which some semantic categories go from being either more frequent than expected to less frequent with the concept, or vice versa. For instance, nouns belonging to EARTH, one of the categories in which the adjectives are used only in sense 1, occur more frequently than expected in P1 and P2, but less frequently in P3 and P4. This is also the case of SENSATION nouns, which are more frequent than expected in P1, but less frequent in the three remaining periods. The reverse holds for nouns belonging to MATTER, since they go from being less frequent than expected in P1 to becoming more frequent in the rest of the periods (cf. Figure 6).

The results discussed in the last paragraph seem to strengthen the interpretation put forward before: whereas the use of the adjectives in sense 1 decreases substantially, their use in sense 2 increases considerably. However, an analysis with a new variable, "sense", would be beneficial and probably necessary to further reinforce this claim. This is so because, as mentioned in Section 3.2, it has been proven that the three semantic categories of nouns SENSATION, MATTER, and BODY include occurrences of the adjectives in both sense 1 and 2. Thus, this further variable would make it possible to tease apart the two senses and determine whether there is in fact a decrease and increase of senses 1 and 2, respectively.

4.4 General discussion of the results

The present discussion focuses on three different sets of results: (1) the concept SWEET-SMELLING, (2) the relationship between the members of the synonym set,

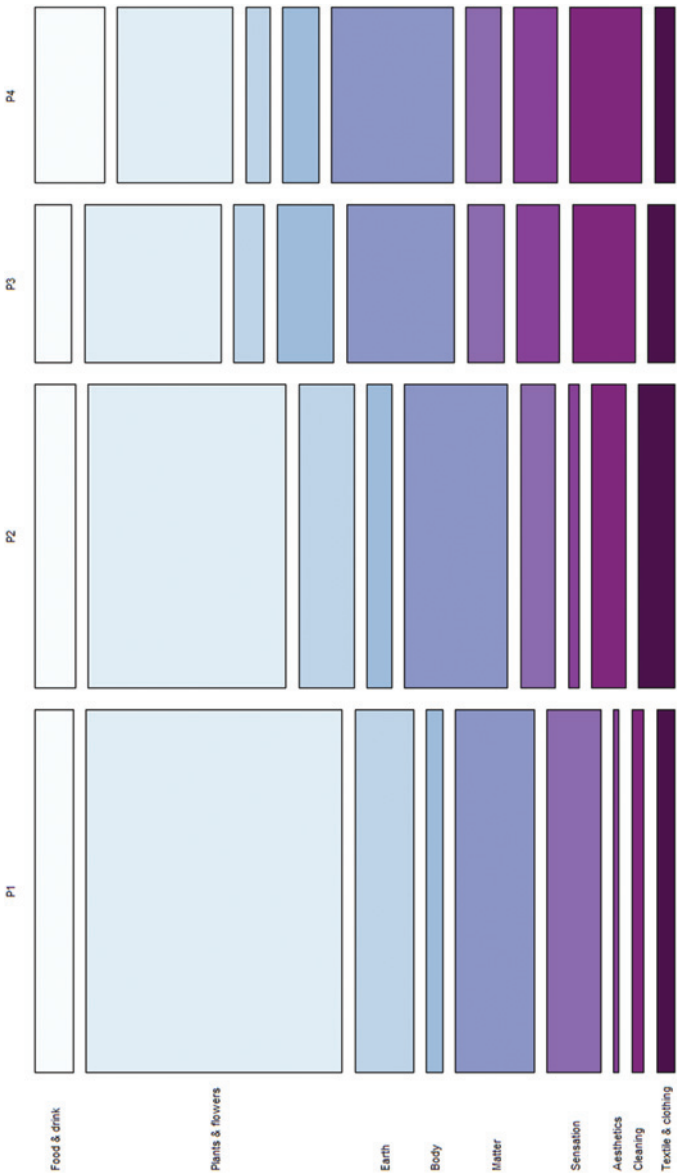

Figure 6: Changes in the distribution of the concept of SWEET-SMELLING across the semantic categories throughout 1850–2009.

and (3) the individual development in collocational behavior of each of the four adjectives. First, as concerns the concept, there is a general decrease in frequency over 1850–2009, which is due mostly to the drastic decline of *fragrant*,

the most frequent adjective in the set, but also partly to the decrease of *perfumed*, which starts out as the second most frequent synonym in P1 but ends up as the second less frequent in P4. Moreover, it has been shown that this decrease mainly takes place with the uses of the adjectives modifying nouns in the semantic category PLANTS & FLOWERS, which decreases drastically over the latter part of LModE and PDE. However, despite this downward tendency in frequency of SWEET-SMELLING, the nouns in CLEANING and AESTHETICS increase significantly, especially in P3 and P4.

Second, with regard to the relation between the four synonyms, it can be noted that *fragrant* and *perfumed* are gradually losing ground in favor of *scented* and, to some extent, *sweet-smelling*, since the last two increase in frequency throughout 1850–2009. In fact, in the case of *scented*, this growth is not that gradual, since it overtakes *perfumed* as the second most frequent adjective already in P2, and by P4 it makes up for 18.1% more than the uses of *perfumed*. Besides, in P4, the frequency of *scented* is not far from that of *fragrant*, which totals 38.9% of the concept, only 4% more than the use of *scented* (34.8%).[11] In conclusion, there seems to be a general gradual convergence of frequency of the four adjectives throughout 1850–2009 (see Figure 1).

Last, concerning the individual development of each of the four adjectives, it is important to highlight that while *perfumed* and *scented* decrease with some categories of nouns (i.e. PLANTS & FLOWERS), both increase with other noun categories, i.e. CLEANING and AESTHETICS (only in the case of *perfumed*). In contrast, *fragrant* only decreases, and this decline is narrowed down to one noun category: PLANTS & FLOWERS. The case of *perfumed* is exceptionally interesting, given that although this adjective suffers a general decline, it still increases with two sorts of nouns, namely AESTHETICS and CLEANING.

The increase in sense 2, both of the concept as well as of some of the individual synonyms (i.e. *perfumed* and *scented*) could possibly be due to the change society has experienced throughout the time span considered in the analysis (1850–2009). First, there has been a growth in the manufacture of products which exhibit an artificially pleasant sweet smell, specifically of those pertaining to the domains of cosmetics, hygiene, and cleansing agents. Additionally, this has resulted in a higher availability of such products and a wider selection of brands. Second, this growth has gone hand in hand with changes in people's habits regarding aspects such as the attention paid to personal hygiene and physical appearance. Third, the advertising industry, given its increasing presence and impact on our choices, has probably also

11 For an interesting method to identify whether particular words take over the functions of others over time, see Baker's (2017: 96–100) discussion on how *on* and *around* are replacing *upon* and *round*, respectively, as the former are drawing away collocates of the latter.

played an important role in making these kinds of products visible and central in our daily lives. This semantic development could thus be understood as the reflection of a changing society, whereby modernization and mass production have led to an ever-increasing need to refer to *artificially* scented soaps and oils rather than to *naturally* fragrant flowers and trees. However, this explanation of the results obtained is yet a tentative one, and further research is still needed to clarify this issue.

5 Conclusions

By conducting a corpus-based analysis and employing the HCFA, this paper has succeeded in establishing the main collocational patterns of the four synonymous adjectives *fragrant, perfumed, scented*, and *sweet-smelling* across the nine semantic categories of nouns over the 160 years analyzed. The findings have both methodological and theoretical implications which pave the way for further diachronic research on the semantic structure of sets of synonyms. This case study adds to the recent literature on the semantic structure of specific sets of near-synonyms. It suggests that innovative methods within the contextual approach to semantics could be applied to the diachronic study of synonyms with equally effective outcomes as those achieved in previous synchronic research. However, to fully assess the generalizability of these methods to diachronic studies on synonymy, they need to be applied in further research to other synonym sets.

By analyzing the attributive uses where the head nouns directly follow the adjectival synonyms, the results provide significant insights about differences in usage between the lexical alternatives, differences which were not fully discernible in the existing descriptions offered by dictionaries and thesauri. Additionally, the diachronic analysis has revealed that the concept SWEET-SMELLING experiences a major change over the time span examined —which might still be under process— since nouns designating a natural smell are modified significantly less frequently by the four synonyms at the end of the period, while the probability of those denoting an artificial smell increases significantly. This change is reflected in the relationships between the members of the synonym set, as they also vary over time, both regarding frequency and collocational patterning. This piece of research therefore provides evidence that both independent variables, "period" and "semantic category of head noun", exert a significant influence on synonym choice, and that there exists a relation between the two. However, these results cannot be generalized

This study is exploratory in nature, but it has still yielded many interesting findings and has proven that diachronic research on synonymy is fruitful and

necessary. However, methodologically speaking, it can still be improved. For instance, including a further variable, "sense", would throw more light on what is going on with the concept SWEET-SMELLING and the lexical items denoting it. Moreover, additional co-occurrence features such as syntactic ones, i.e. including the "syntactic function of the adjectives" (e.g. predicative, attributive, and postponed), or stylistic ones (e.g. "text-type") would indeed benefit the interpretations of the results put forward in the discussion. For example, by analyzing the text-types in which the adjectives are used in sense 2, it might be possible to confirm if the rise of this sense could reflect a change to a modernized society characterized by mass-production and commercialism. For instance, it would be beneficial to examine the genre magazine and news in more detail, as a growth in these text-types could help determining whether the hypothesis put forward in the discussion is true. Additionally, different periodizations could be tested, for instance, to see whether changes in frequency between the synonyms and senses are related to periods of economic growth/decline. The analysis could also be conducted with more elaborate statistical tools such as regression models or conditional inference trees, as well as visualization tools such as Multidimensional Scaling put forward in recent diachronic studies of the same type but on other semantic phenomenon such as polysemy (Jansegers and Gries 2017). Therefore, much remains to be done in historical cognitive semantics, specifically in the domain of synonymy, which constitutes a field that will probably prompt much research in the near-future.

Acknowledgements: This research was conducted with the financial support of the European Regional Development Fund and the following institutions: Regional Government of Galicia (grants ED481A-2016/168, ED431D 2017/09, and ED431B 2017/12) and the Spanish Ministry of Economy and Competitiveness (grant FFI2017-86884-P). I am indebted to María José López-Couso, Dirk Speelman, and Dirk Geeraerts for insightful discussions of the results. I am also grateful to Iván Tamaredo and Mario Serrano-Losada for their most helpful suggestions on earlier versions of this paper. Finally, thanks are also due to two anonymous reviewers and the editor of *Corpus Linguistics and Linguistic Theory*, Stephanie Wulff, for their time and consideration, as well as for their fruitful comments.

References

The American heritage dictionary of the English language. 2016–. Houghton Mifflin Harcourt. Online version: https://ahdictionary.com/ (accessed 04 March 2018).

Baker, Paul. 2017. *American and British English: Divided by a common language?* Cambridge: Cambridge University Press.
Berlin, Brent & Paul Kay. 1969. *Basic color terms: Their universality and evolution*. Berkely & Los Angeles: University of California Press.
Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. (Cambridge Approaches to Linguistics). Cambridge: Cambridge University Press.
Biggam, Carol. P. 2012. *The semantics of colour. A historical approach*. Cambridge: Cambridge University Press.
Cambridge Dictionary. 2016–. Cambridge University Press. Online version: https://dictionary.cambridge.org/ accessed 04 March 2018.
Church, Kenneth Ward, William Gale, Donald Hindle & Rosamund Moon. 1994. Lexical substitutability. In Beth Levin & Antonio Zampolli (eds.), *Computational approaches to the lexicon*, vol. 255, 153–177. Oxford & New York: Oxford University Press.
Church, Kenneth Ward & Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16(1). 76–83.
Collins Dictionary. 2012–. Collins. Online version: https://www.collinsdictionary.com/ (accessed 04 March 2018).
Cruse, D. Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
Davies, Mark. 2008. *The Contemporary Corpus of American English*. https://corpus.byu.edu/coca/.
Davies, Mark. 2010. *The Corpus of Historical American English*. https://corpus.byu.edu/coha/.
Divjak, Dagmar. 2006. Ways of intending: Delineating and structuring near-synonyms. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in Cognitive Linguistics*, 19–56. Berlin & New York: Mouton de Gruyter.
Divjak, Dagmar & Stefan Th Gries. 2006. Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2(1). 23–60.
Edmonds, Philip & Graeme Hirst. 2002. Near-Synonymy and Lexical Choice. *Computational Linguistics* 28(2). 105–144.
English Oxford Dictionaries. 2016–. Oxford University Press. Online version: https://www.oxforddictionaries.com/ (accessed 04 March 2018).
Firth, J. R.. 1957a. *Studies in linguistic analysis*. Oxford: Blackwell.
Firth, J. R.. 1957b. *Papers in linguistics, 1934-1951*. London & New York: Oxford University Press.
Geeraerts, Dirk. 1986. On necessary and sufficient conditions. *Journal of Semantics* 5(4). 275–291.
Geeraerts, Dirk. 2010. *Theories of lexical semantics*. Oxford: Oxford University Press.
Gries, Stefan Th. 2001. A corpus-linguistic analysis of English *-ic* vs *-ical* adjectives. *ICAME Journal* 25(i). 65–108.
Gries, Stefan Th. 2003. Testing the sub-test: An analysis of English *-ic* and *-ical* adjectives. *International Journal of Corpus Linguistics* 8(1). 31–61.
Gries, Stefan Th. 2004. *HCFA 3.2. A Program for R*. http://www.linguistics.ucsb.edu/faculty/stgries.
Gries, Stefan Th. 2009. *Statistic for linguistics with R*. Berlin: Mouton de Gruyter.
Gries, Stefan Th. 2010. Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon* 5(3). 323–346.
Gries, Stefan Th & Dagmar Divjak. 2009. Behavioral profiles: A corpus-based approach to cognitive semantic analysis. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in cognitive linguistics*, 57–75. Amsterdam: John Benjamins.

Gries, Stefan Th & Naoki Otani. 2010. Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal* 34. 121–150.
Henning, Hans. 1916. *Der Geruch*. Leipzig: JA Barth.
Heylen, Kris, Yves Peirsman, Dirk Geeraerts & Dirk Speelman. 2008. Modelling word similarity: An evaluation of automatic synonymy extraction algorithms. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC)*, 3243–3249. Marrakech: European Language Resources Association.
Hilpert, Martin. 2013. *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
Jansegers, Marlies & Stefan Th Gries. 2017. Towards a dynamic behavioral profile: A diachronic study of polysemous *sentir* in Spanish. *Corpus linguistics and linguistic theory*. doi: 10.1515/cllt-2016-0080 (accessed 07 January 2018).
Kay, Christian & Kathryn Allan. 2015. *English historical semantics*. Edinburgh: Edinburgh University Press.
Kennedy, Graeme. 1991. *Between* and *through*: The company they keep and the function they serve. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 95–110. London: Longman.
Kjellmer, Göran. 1995. Synonymy and corpus work: On *almost* and *nearly*. *ICAME Journal* 27. 19–27.
Lehrer, Adrienne. 1969. Semantic cuisine. *Journal of Linguistics* 5. 39–55.
Lehrer, Adrienne. 1974. *Semantic fields and lexical structure*. Amsterdam & London: North Holland.
Levshina, Natalia. 2011. *Doe wat je niet laten kan* (A usage-based analysis of Dutch causative constructions). Leuven: KU Leuven.
Levshina, Natalia. 2015. *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam & Philadelphia: John Benjamins.
Levshina, Natalia & Kris Heylen. 2014. A radically data-driven construction Grammar: Experiments with Dutch causative constructions. In Ronny Boogaart, Timothy Colleman & Gijsbert Rutten (eds.), *Extending the scope of construction grammar* (Cognitive Linguistic Research), 17–46. Berlin: Mouton de Gruyter.
Liu, Dilin. 2010. Is it a *chief, main, major, primary*, or *principal concern*? A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics* 15(1). 56–87.
Liu, Dilin. 2013. Salience and construal in the use of synonymy: A study of two sets of near-synonymous nouns. *Cognitive Linguistics* 24(1). 67–113.
Liu, Dilin & Maggie Espino. 2012. *Actually, Genuinely, Really*, and *Truly*: A corpus-based Behavioral Profile study of near-synonymous adverbs. *International Journal of Corpus Linguistics* 17(2). 198–228.
Longman dictionary of contemporary English. 2015–. Pearson. Online version: https://www.ldoceonline.com/ (accessed 04 March 2018).
Lorig, Tyler S. 1999. On the similarity of odor and language perception. *Neuroscience & Biobehavioral Reviews* 23. 391–398.
McEnery, Tony, Richard Xiao & Yukio Tono. 2006. *Corpus-based language studies: An advanced resource book*. Abingdon and New York: Routledge.
Merriam-webster dictionary and thesaurus. 2017–. Merriam-Webster. Online version: https://www.merriam-webster.com/ (accessed 04 March 2018).
Murphy, M. Lynne. 2003. *Semantic relations and the lexicon: Antonymy, synonymy and other paradigms*. Cambridge; New York: Cambridge University Press.

Oxford English Dictionary, 3rd ed. 2000–. Oxford University Press. Online version: http://www.oed.com/ accessed 04 March 2018.
Partington, Alan. 1998. *Patterns and meanings: Using corpora for English language research and teaching*. (Studies in Corpus Linguistics). Amsterdam & Philadelphia: John Benjamins.
R Core Development Team. 2015. R: A Language and Environment for Statistical Computing. Viena: The R Foundation for Statistical Computing. http://www.R-project.org.
Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Translated by Roy Harris as *Course in general linguistics*.1983. London: Duckworth.
Sinclair, John, *et al*. 1966. Beginning the study of lexis. In Charles E Bazell (ed.), *In memory of J.R. Firth*, 410–429. Harlow: Longman.
Sinclair, John. 1987. Collocation: A progress report. In Ross Steele & Terry Threadgold (eds.), *Language topics: Essays in honor of Michael Halliday*. Amsterdam & Philadelphia: John Benjamins.
Sinclair, John. 2004. *Trust the text: Language, corpus and discourse*. Monograph. London: Routledge.
Speelman, Dirk & Dirk Geeraerts. 2009. Causes for causatives: The case of Dutch *doen* and *laten*. In Ted Sanders & Eve Sweetser (eds.), *Causal categories in discourse and cognition*, 173–204. Berlin & New York: Mouton de Gruyter.
Stubbs, Michael. 2001. *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.
Taylor, John R. 2003. Near synonyms as co-extensive categories: "High" and "tall" revisited. *Language Sciences* 25(3). 263–284.
Yeshurun, Yara & Sobel Noam. 2010. An odor is not worth a thousand words: From multidimensional odors to unidimensional odor objects. *Annual Review of Psychology* 61. 219–241.

Bionote

Daniela Pettersson-Traba

Daniela Pettersson-Traba holds a BA in English Language and Literature (July 2014) and an MA in English Studies (September 2015). She is currently a full-time postgraduate researcher at the Department of English and German of the University of Santiago de Compostela (Spain), under funding from the Regional Government of Galicia (grant ref. ED481A-2016/168).