

Analyzing the Behavioral Profiles of Sets of Near-synonyms in American English from a Diachronic Perspective

D. Pettersson-Traba¹

¹ Universidad de Santiago de Compostela, Departamento de Filología Inglesa y Alemana de la Facultad de Filología de la Universidad de Santiago de Compostela, E-15782 Santiago de Compostela - España, danielatraba@hotmail.com

1. INTRODUCTION

It is a rather generalized assumption that synonymy is relatively straightforward and unproblematic, being the semantic relation which is familiar to most people, including non-linguists. However, despite being a common linguistic phenomenon, synonymy is also a particularly complex one (Cruse, 2000; Liu, 2010). This can be demonstrated by the fact that choosing the most suitable word from a set of potential synonyms in particular contexts of use is not considered to be an easy task, but has been and continues to be a considerable “significant language generation problem” (Gardiner & Dras, 2007: 31). The reason for this is that the vast majority of synonyms existing in languages are not absolute synonyms but near-synonyms, and therefore entail a certain degree of contrast. As such, synonyms are defined as semantically related words which share the same denotational core meaning, but which differ in peripheral aspects or in other dimensions of meaning such as style, connotation, or collocation¹ (Edmonds & Hirst, 2002; Murphy, 2003). Thus, synonyms can sometimes be used interchangeably, though not always. In other words, synonyms “are neither in free variation, nor in complementary distribution” (Divjak, 2006: 21). Consider the following examples with the near-synonyms *prize* and *award*:

- (1) What’s a synonym for *prize*? – *Award*
- (2) The plaintiff received a hefty *award* (*≠ prize*) in the lawsuit.
- (3) Jan won the *prize/award* for the best drawing.

(from Murphy, 2003: 137)

Both *prize* and *award* designate a ‘thing given as recognition of outstanding achievement’ and hence share the same core denotational meaning. Consequently, they are considered synonymous in a neutral context such as (1). However, despite their obvious semantic similarities, they differ as regards peripheral aspects of their denotational meaning, a result of their different senses and extensions. This makes them interchangeable in some non-neutral contexts such as the one in (3), whereas they are not similar enough to be freely interchangeable in others such as that in (2).

This example clearly demonstrates that knowledge about synonymic differences is crucial to understand how sets of synonyms work in terms of nuances of meaning and usage patterns, although, interestingly, synonyms are usually defined only in terms of their similarities. While most existing synonym descriptions in lexicographical resources offer valuable information about the semantic attributes lexical items share, they do not provide

¹ Throughout this paper I will employ the terms ‘synonym’/‘synonymy’ and ‘near-synonym’/‘near-synonymy’ interchangeably, since many linguists (Cruse, 2000; Liu & Espino, 2012, Murphy, 2003, among others) concur that absolute or true synonymy is very infrequent in languages, if it exists at all.

a comprehensive view of their differences. Thus, these reductive descriptions prevent users of dictionaries and thesauri from fully comprehending how particular synonyms are to be differentiated and in which contexts they can be freely interchangeable (Liu, 2010; Liu & Espino, 2012). However, it should be noted that in order to provide a complete picture of how groups of synonyms differ as regards meaning and usage patterns, it is necessary to uncover their internal semantic structure. This can only be done by accounting for the dimensions of meaning such as style and collocation, which tend to be overlooked or forgotten.

2. THE CORPUS-BASED BEHAVIORAL PROFILE (BP) APPROACH: THEORETICAL FRAMEWORK AND METHODOLOGICAL ISSUES

A considerable number of the existing corpus-based studies have focused on the areas of lexis and semantics (e.g. polysemy, idioms, phrasal verbs). The corpus-based behavioral profile (henceforth BP) approach has proven to be particularly effective in these domains. The underlying assumption of this approach is that the meanings of lexical items cannot be fully understood without taking into account their BPs, i.e. their distributional patterns, be they collocational, syntactic, or stylistic (Divjak, 2006; Gries & Otani, 2010; Liu, 2010). Nonetheless, in spite of its usefulness for accurate and comprehensive descriptions of lexico-semantic usage, only a few have focused on synonyms. As Divjak (2006) points out, the internal semantic structure and BPs of particular groups of synonyms have “hitherto remained largely undiscussed in the literature” (33). This suggests that there exists a significant gap in this field of research that needs to be filled in order to gain a better understanding of how specific synonyms differ.

Regardless of the limited number of corpus-based BP studies dealing with synonymy, the few which have been carried out on sets of near-synonymous verbs (e.g. Hanks, 1996; Divjak, 2006; Divjak & Gries, 2006), adjectives (Liu, 2010; Gries & Otani, 2010), and adverbs (Liu & Espino, 2012) have proven this approach to be particularly valuable for delineating fine-grained aspects of meaning and for unfolding differences among synonyms. For instance, Liu’s (2010) study on the near-synonyms *chief*, *main*, *major*, *primary*, and *principal* identified significant usage patterns among the five adjectives by analyzing their collocational, colligational², and stylistic behavior. Liu paid special attention to the types of nouns each adjective modifies, and by doing so affirmed that analyzing the head nouns of adjectives, in particular attributive adjectives, is an effective way of capturing the “essence of the semantics of adjectives” (2010:56). Additionally, some of Liu’s results challenged previous descriptions of the synonyms. Liu found out that the information included in the reference materials is too limited to distinguish among them. Likewise, he discovered that the definitions and examples of usage are sometimes inaccurate, as they do not reflect how the adjectives are employed in actual language use.

This very brief review of the BP approach demonstrates its effectiveness in identifying important fine-grained semantic and usage differences among near-synonyms. By taking into account a wider range of contextual and distributional characteristics than most other

² Colligations are the collocation of a node word with a particular grammatical class of words. Liu analyzed two types of colligational behavior: the frequency of the selected adjectival synonyms with (i) the singular and plural tokens of its noun collocates and (ii) indefinite and definite determiners.

corpus-based studies, the BP approach offers a comprehensive picture of the usage patterns and semantic structure of lexical items. Nevertheless, much remains to be done to clarify how individual groups of synonyms work in terms of usage patterns and nuances of meaning.

Against this backdrop, my PhD thesis proposes a diachronic corpus-based BP analysis of sets of adjectival near-synonyms of the type *crucial*, *essential*, *indispensable*, and *vital* or *amusing*, *comical*, *funny*, and *humorous*. The principal objective is to discover their usage patterns and fine-grained aspects of meaning by paying special attention to their internal semantic structure. This is done by examining and comparing the contextual and stylistic environments in which the synonyms are used, thus aiming at unveiling differences in behavior among them, not only regarding semantics, but also concerning syntax and style. Throughout the analysis I also intend to establish the co-occurrences which best help to reveal the nature of the semantics of the adjectives at issue. In addition, this study tries to identify the main usage patterns of the adjectives in different periods of the recent history of American English so as to trace the diachronic evolution of their BPs. The linguistic data which will be used is extracted from the 400 million-word computerized diachronic corpus the *Corpus of Historical American English (COHA)*, which covers the period 1810-2009.

All in all, by employing the corpus-based BP approach, the intention of this thesis is to add to the existing yet scarce literature on synonymy and to test the applicability and validity of the BP approach on diachronic data, thus examining the internal semantic structure of particular groups of near-synonyms from a diachronic perspective, something which no previous BP study has done.

3. A PILOT STUDY: SOME PRELIMINARY RESULTS

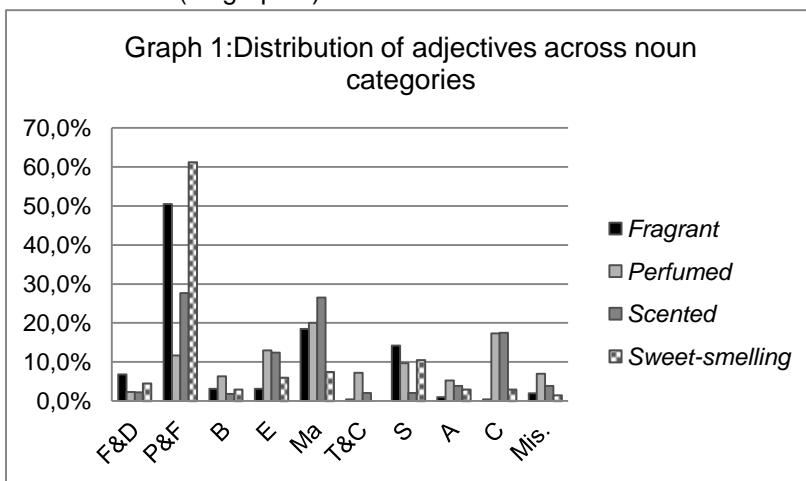
My MA dissertation served as a point of departure for my PhD thesis. The principal objective of this diachronic exploratory BP study was to examine the collocational behavior of the attributive uses of the synonymous adjectives *fragrant*, *perfumed*, *scented*, and *sweet-smelling* by focusing on the nouns they typically modify (Pettersson-Traba, 2015).

This pilot study succeeded in establishing the main usage patterns and distributions of the four adjectives across different categories of nouns, thus revealing important differences in contextual behavior among them. For instance, nouns categorized into 'cleaning' (e.g. SOAP and HANDKERCHIEF) and 'the earth' (e.g. SEA and WATER) are modified mainly by *perfumed* and *scented*, whereas 'sensation' nouns (e.g. SMELL and AROMA) are mostly modified by *fragrant*. Only one category, namely 'plants and flowers' (e.g. BLOSSOM and GRASS) are modified fairly often by the four adjectives (cf. graph 1)³. However, a detailed examination showed that the adjectives co-occur with different sub-categories, and only one noun, FLOWER, is frequently modified by all four adjectives.

Moreover, the results show that some adjectives in the set are more similar than others. For instance, *perfumed* and *scented* are the only adjectives which denote an 'artificial sweet and pleasant smell', since nouns referring to man-made objects (e.g. SOAP, GARMENT, and CREAM) are modified mainly by them. To illustrate this point, in the 'cleaning'

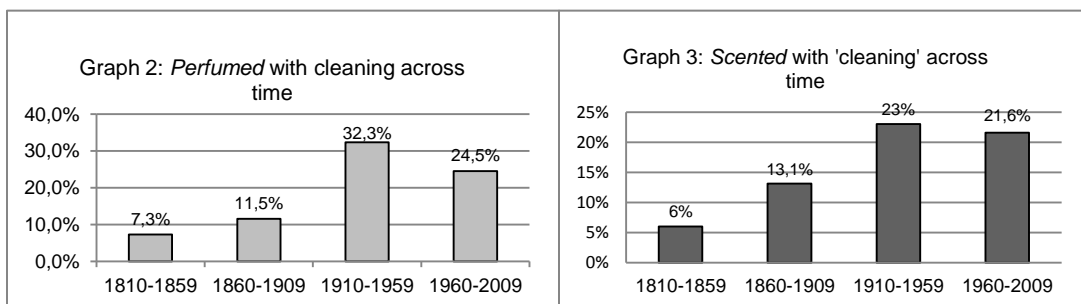
³ The abbreviations used in this graph stand for the following: F&D = food and drink, P&F = plants and flowers, B = the body, E = the earth, Ma = matter, T&C = textile and clothing, S = sensation, A = aesthetics, C = cleaning, and Mis. = Miscellaneous.

category, there are only 3 (0.4 %) and 2 (3%) examples of *fragrant* and *sweet-smelling*, respectively, whereas a total of 52 (17.3 %) and 58 (17.5 %) examples contain *perfumed* and *scented* as modifiers (cf. graph 1).



All things considered, the results obtained in this part of the analysis show that despite seeming nearly identical in the lexicographical resources, in which the four adjectives share the same basic and central definition, ‘having a sweet and pleasant smell’, they have different collocational affinities since they tend to collocate with different types of nouns. This corroborates findings obtained in previous BP studies, namely that dictionary descriptions of synonyms are too limited to gain an understanding of how they differ, and hence makes it difficult to know in which contexts synonyms can be used interchangeably.

Moreover, the diachronic analysis of each adjective yielded some revealing changes in usage patterns. For example, nouns grouped under ‘plants and flowers’ have undergone a general substantial decline. Also, nouns belonging to the category ‘cleaning’ are more frequently modified by *perfumed* and *scented* in the period 1910-2009 than in 1810-1909. This seems to point to an increase in their use when employed to denote an ‘artificial sweet and pleasant smell’ (cf. graphs 2 and 3). In short, it seems that this study, although exploratory in nature, has confirmed the applicability and usefulness of the BP approach for the diachronic analysis of near-synonyms.



In view of the results of this pilot study, my PhD offers a more exhaustive diachronic BP analysis of sets of adjectival synonyms. To this purpose, I intend to examine a larger number of usage patterns. Besides analyzing the collocational behavior, I also take the colligational, syntactic, and stylistic patterns of the selected synonyms into consideration.

4. Conclusions

Although synonymy is far from being a trivial problem, research conducted on this semantic relation is scarce. In particular, the internal semantic structure of particular groups of synonyms has often been left aside in the specialized literature. In spite of this, the few corpus-based BP studies which have concentrated on individual sets of synonyms have proven this approach to be especially effective in providing an understanding of how synonyms work in terms of fine-grained aspects of meaning and usage patterns.

However, further research is necessary to provide a more accurate and satisfactory description of this semantic relation. This study will, therefore, partially cover an existing gap within this field by analyzing the BPs of sets of adjectival near-synonyms from a diachronic perspective. The intention is to discover their nuances of meanings and usage patterns, to compare their BPs, and to trace their semantic development over the last couple of centuries in American English. The main aims are to (i) further test the validity of the BP approach when applied to adjectival synonyms and (ii) to test and probe its applicability and usefulness for the diachronic analysis of synonyms.

References:

- [1] Cruse, A. D. (2000). *Meaning in language: An introduction to semantics and pragmatics*. Oxford: O. U. P.
- [2] Divjak, D. (2006). Ways of intending: Delineating and structuring near synonyms. In S. Th. Gries & A. Stefanowitsch (Eds.), *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis* (pp. 9-56). Berlin & New York: Mouton de Gruyter.
- [3] Divjak, D., & St. H. Gries. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2 (1), 23-60. doi: 10.1515/cllt.2006.002
- [4] Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics*, 28(2), 105-144. Retrieved from <http://www.mitpressjournals.org/loi/coli>
- [5] Gardiner, M., & M. Dras. (2007). Exploring approaches to discriminating among near-synonyms. *Proceedings of the Australasian Language Technology Workshop*, 31-39. Retrieved from <http://www.aclweb.org/anthology/U07-1007>
- [6] Gries, S. Th., & N. Otani. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal* 34, 121-150. Retrieved from <http://clu.uni.no/icame/>
- [7] Hanks, P. (1996). Contextual dependency and lexical sets. *International Journal of Corpus Linguistics* 1(1), 75-98. doi: 10.1075/ijcl.1.1.06han
- [8] Liu, D. (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. doi: 10.1075/ijcl.15.1.03liu
- [9] Liu, D., & M. Espino. (2012). *Actually, genuinely, really, and truly*. A corpus-based behavioral profile study of the near-synonymous adverbs. *International Journal of Corpus Linguistics* 17, 198-228. doi: 10.1076/ijcl.17.2.03liu
- [10] Murphy, L. M. (2003). *Semantic relations and the lexicon: Antonymy, synonymy, and other paradigms*. Cambridge & New York: C.U.P.
- [11] Pettersson-Traba, D. (2015). A preliminary corpus-based diachronic analysis of the behavioral profile of a set of near-synonyms in American English. Unpublished MA thesis, University of Santiago de Compostela.