

# Use of Hedges in Definitions: Out of Necessity or Theory-Driven?

Mariusz Piotr Kamiński, *University of Applied Sciences, Nysa,  
Poland (mariusz.kaminski@pwsz.nysa.pl)*

---

**Abstract:** Language has an inventory of words and expressions (e.g. *especially, sort of, loosely speaking*) used to communicate that what is being said is not exactly precise or complete. Referred to as hedges, they provide support for the conception of prototypically organized categories, developed by Eleanor Rosch in the 1970s and elaborated by her followers in subsequent decades. Given the fact that hedges are extremely useful for lexicographers in defining, this paper examined the frequency and distribution of hedges in major English dictionaries over the past centuries. One of the findings of this research is the fact that, although hedges have been used in English lexicography since Johnson-1785, their use has intensified in certain dictionaries since the rise of prototype theory, suggesting that recent defining practice in these dictionaries must have been influenced by this theory. Other factors determining the use of hedges were explored.

**Keywords:** HEDGES, DICTIONARIES, LEXICOGRAPHY, DEFINITION, DEFINITION LANGUAGE, PROTOTYPE THEORY, HISTORY, FREQUENCY, DISTRIBUTION, DEFINING STYLE

**Opsomming: Die gebruik van vaaghede in definisies: Noodsaaklik of teoriegedrewe?** Taal beskik oor 'n lys woorde en uitdrukkings (bv. *veral, 'n soort, oor die algemeen*) wat gebruik word om aan te dui dat wat gesê word nie heeltemal akkuraat of volledig is nie. Hierdie terme, wat vaaghede genoem word, verleen steun aan die voorstelling van prototipies georganiseerde kategorieë wat in die 1970's deur Eleanor Rosch ontwikkel is en deur haar volgelinge in daaropvolgende dekades verfyn is. Aangesien vaaghede uiters nuttig vir leksikograwe is tydens definiëring, is die frekwensie en verspreiding daarvan in die belangrikste Engelse woordeboeke van die afgelope eeue in hierdie artikel ondersoek. Een van die bevindings van hierdie navorsing is dat, alhoewel vaaghede sedert Johnson-1785 in Engelse leksikografie gebruik is, hul gebruik in sekere woordeboeke toegeneem het sedert die opkoms van die prototipeteorie, wat daarop dui dat onlangse definieerpraktyk(e) in hierdie woordeboeke deur dié teorie beïnvloed moes gewees het. Ander faktore wat die gebruik van vaaghede bepaal, is ook verken

**Slutelwoorde:** VAAGHEDE, WOORDEBOEKE, LEKSIKOGRAFIE, DEFINISIE, DEFINISIE-TAAL, PROTOTIPETEORIE, GESKIEDENIS, FREKWENSIE, VERSPREIDING, DEFINIEERSTYLE

## 1. Introduction

Language has an inventory of words and expressions used to communicate that what is being said is not exactly precise or complete. George Lakoff has referred to such metalinguistic devices as hedges, explaining that their "mean-

ing implicitly involves fuzziness" and their "job is to make things fuzzier or less fuzzy" (1973: 471). From a syntactic point of view, a great number of hedges are adverbs and sentence adverbials which function in sentences as modifiers, for example: *loosely speaking*, *strictly speaking*, *technically*, *especially*, *chiefly*, *specifically*, *in particular*, *very*. However, hedges are not a grammatically uniform group, as they include other word categories and even punctuation marks.<sup>1</sup> As Ken Hyland (1994 and 2006) remarks, a hedge is any metalinguistic device that marks uncertainty, hesitation, ambiguity, and tentativeness.

Hedges express varying degrees of category membership. As Lakoff demonstrates, a hedge *par excellence* indicates the most central member of a category (as in *A robin is a bird par excellence*), while *sort of* points to a peripheral example of the category (as in *A penguin is sort of a bird*). What is more, *loosely speaking* does not merely exclude the central member but points to "things that would not ordinarily be considered members" (as in *Loosely speaking, a telephone is a piece of furniture*) (Lakoff 1973, Taylor 1995: 77). Such hedges as *loosely speaking* show that not only are category boundaries flexible, but they can also be redefined. These findings provide support for the conception of prototypically organized categories, whereby categories display degrees of typicality and their boundaries are blurred (Rosch 1973; Lewandowska-Tomaszczyk 2007: 145).

What bearings do the above findings have on defining practice? From the point of view of practical lexicography, hedges are convenient tools that make the definer's task much easier. Rather than describing all possible exemplars and features of the category of meaning being defined, which is often doomed to failure, the definer highlights what is typical of this category.<sup>2</sup> The former situation reflects the classical approach to meaning, which had dominated linguistic thought since Aristotle. This approach rests on the assumption that a meaning is identified through a fixed set of necessary and sufficient features. The same assumption underpins the construction of the traditional definition (*genus proximum plus differentia specifica*),<sup>3</sup> which presupposes that the lexicon is organized hierarchically and that lexemes can be defined in discrete rather than continuous terms.<sup>4</sup> If lexicographers adhere uncritically to the classical principle of defining, they end up with verbose and over-specific definitions<sup>5</sup> (Atkins and Rundell 2008). In the attempt to search for necessary conditions for the use of a word, lexicographers are likely to obscure central meanings by marginal ones. Hedges do not merely emphasize the prototype but downtone potentially controversial features; for example, as Geeraerts (2006) points out, fruit is usually sweet, not sour like lemons.

Prototype theory is relatively new in the history of linguistics. It was developed by a cognitive psychologist Eleanor Rosch in the 1970s through a series of experiments on the internal structure of categories. Empirical evidence collected by Rosch and her colleagues attracted considerable attention of linguists. Since the early 1980s the theory has exerted a great impact on cognitive linguistics, which is reflected by the growing body of literature with prototypes

in the center of scholarly attention (e.g. Lakoff 1987; Langacker 1987; Taylor 1995; Geeraerts 1997; Lewandowska-Tomaszczyk 2007).

Prototype theory has recently drawn a great deal of attention from dictionary makers and researchers (Hanks 1994, Rey 1990, Swanepoel 1994, Van der Meer 1999 and 2000, Geeraerts 2006, Zgusta 2006). In a discussion of the relationship between prototypicality and lexicography, Geeraerts concludes that cognitive semantics offers "an exciting perspective for the further development of lexicography" (Geeraerts 2006: 363). It is also Zgusta (2006: 115) who notes that "prototype theory promises to be of great usefulness to lexicography". The *New Oxford Dictionary of English* (NODE-1998) is perhaps one of the first dictionaries to refer explicitly to cognitive linguistics as a basis for the description and presentation of meaning (cf. Geeraerts 2006, Hanks 1994). In the Preface, we read:

The *New Oxford Dictionary of English* is [...] informed by currently available evidence and current thinking about language and cognition. [...] Linguists, cognitive scientists, and others have been developing new techniques for analysing usage and meaning, and the *New Oxford Dictionary of English* has taken full advantage of these developments. Foremost among them is an emphasis on identifying what is 'central and typical' [...] The style of definition adopted for the *New Oxford Dictionary of English* aims in part to account for the dynamism, imaginativeness, and flexibility of ordinary usage. [...] The layout and organization of each entry in the dictionary reflect this new approach to meaning. Each entry has at least one core meaning, to which a number of subsenses, logically connected to it, may be attached. (Preface, NODE-1998)

In accordance with the prototypical conception of meaning, the NODE-1998 editor recognizes the fact that the semantic structure of a lexical unit is fuzzy and flexible, with clusters of meanings related to one or more prototypical cores. One may expect that lexicographers who adopt the radial structure model will use special techniques for highlighting prototypical elements and for indicating relationships between meanings. One of these techniques is the use of hedges.

For anyone familiar with the history of dictionaries, it is easy to see that hedges are by no means a novelty for contemporary dictionary writers and that they have been used in definitions long before the rise of prototype theory. However, considering the usefulness of prototype theory for lexicographers, it is likely that hedges are applied more consciously now than before the rise of this theory. It follows that the use of hedges might have intensified over the recent decades. This is the main research question that this paper attempts to address. Other points of interest are whether there are other factors determining the use of hedges (besides the rise of prototype theory), and whether English dictionaries display preferences for particular sets of hedges. This paper aims to examine the frequency and distribution of hedges in definitions in major English dictionaries published over the past centuries.

## 2. Method

### 2.1 Materials

The dictionaries under study were published between 1785 and 2011, a period of time covering a large part of the history of English monolingual lexicography:

- Johnson's *Dictionary of the English Language* (henceforth Johnson-1785);
- Webster's *American Dictionary of the English Language* (Webster-1865);
- *The Oxford English Dictionary on Historical Principles* (1888–1928) (OED-1928);
- *Chambers's Twentieth Century Dictionary* (Chambers-1952);
- *The Chambers Dictionary* (Chambers-2011);
- *Webster's Seventh New Collegiate Dictionary* (Collegiate-1963);
- *Merriam-Webster's Collegiate Dictionary* (Collegiate-2004);
- *The New Oxford Dictionary of English* (NODE-1998);
- *The Longman Dictionary of Contemporary English* (LDOCE-2005);
- *The Concise Oxford English Dictionary* (COD-2011)

An advantage of the above selection is that it includes the editions of the same dictionaries: two editions of Chambers and two editions of the Collegiate, published before and after the rise of prototype theory. This material was particularly worth exploring as it ensured the isolation of the effect of publication time on the use of hedges from other variables, notably dictionary type.

Another advantage of the above selection is that it represents the main lexicographic genres that emerged over the past three centuries: general-purpose dictionaries for native speakers, scholarly historical dictionaries for the same readership, and dictionaries for learners. Johnson-1785 was one of the first English dictionaries which aimed at comprehensiveness of vocabulary coverage and thoroughness of treatment. Unlike earlier dictionaries of the 17th century, it aimed to be a scholarly record of the entire language. Along with rare and obsolete words used in English literature, the dictionary covered ordinary everyday words, such as *make*, *do*, *book*, and their meanings. Johnson had a strong influence on subsequent dictionaries both in the United States and Britain, notably Webster-1864 and OED-1928 (Friend 1967; Hanks 2005; Landau 2009), which are also included in this study. Webster aimed to surpass Johnson by covering numerous terms of developing science, art, and technology. His definitions were more exhaustive than Johnson's, with a more elaborate sense division (Landau 2001: 70). OED-1928 is the only historical work in this selection, designed to serve students and scholars as a scholarly resource for research into the history of the English language. NODE-1998, COD-2011, the two editions of the Collegiate (1963 and 2004), and the two editions of Chambers (1952 and 2011), are relatively recent publications for a popular audience. They are single volume works designed to satisfy the needs of a wide range of native speakers especially when they arise while reading. As a dictionary for EFL

learners, LDOCE-2005 represents a relatively new genre in English lexicography. A distinguishing feature of this dictionary is that it uses in its definitions a restricted set of defining vocabulary of around 2000 words. Although the above selection is modest, it covers the dictionaries that have been extremely influential in both English (Johnson-1785, OED-1928) and American lexicography (Webster-1865). The dictionaries under study either shaped English lexicography or grew directly out of the long tradition of dictionary making. They vary considerably one from another with respect to the target audience, the publication date, the amount and type of information provided. All these factors may have an effect on the use of hedges.

## 2.2 Data preparation

Approximately 30 to 60 pages were drawn randomly from each dictionary.<sup>6</sup> The page images were converted into text files using OCR software.<sup>7</sup> The data were further processed by extracting definitions, and removing other parts of entries. The text was proofread and checked. To facilitate the identification of hedges, abbreviated forms such as *esp* and *usu*, which were conventionally used in some of the dictionaries, were expanded to their full forms. For the same purpose, some abbreviations were converted to their variant forms; for example *&c* into *etc.*, and *e.g.* into *eg.* The samples ranged in size from over 16,000 word tokens in the OED-1928, Johnson-1785, and LDOCE-2005 to 37,000 in the Collegiate-1963 (see Table 1), but in order to minimize the effect of different sample lengths, the hedge frequencies were normalized by 100,000. The comparison of two editions of the Collegiate seems to have been least affected by the size differences, as the samples were roughly comparable in length (37664 word tokens in Collegiate-1963 and 35019 in Collegiate-2004). A similar situation obtained in Chambers (23550 in Chambers-1952 and 28153 in Chambers-2011).

83 hedges were selected for the study, with more than half of them (48) being drawn from Lakoff (1973: 472).<sup>8</sup> Due to the fact that Lakoff's list excludes a number of items that typically occur in definitions (e.g. *etc.*, *e.g.*, *also*, *or*, *such as*, *usually*), while including those that are unlikely to be found in this type of text (e.g. *in a manner of speaking*, *loosely speaking*, *mutatis mutandis*), the selection was expanded by 35 additional hedges. These were identified by perusing dictionaries, including those under consideration. Given the fact that it was often impossible to make arbitrary decisions regarding whether a hedge is representative of the defining language, the majority of Lakoff's hedges were retained,<sup>9</sup> regardless of whether they were characteristic of definition style or not. The selection under study is by no means complete (see their grammatical variation in the Introduction), but it includes a large number of hedges that typically occur in definitions. The list is as follows:

"a real", "a regular", "a true", "actually", "all but", "all but a", "almost", "also", "anything but a", "approximately", "as", "as if", "as it were", "assumed", "basi-

cally", "believed", "broadly", "can be viewed as", "chiefly", "considered", "definitely", "eg", "especially", "essentially", "etc", "exceptionally", "for example", "for the most part", "generally", "in a manner of speaking", "in a real sense", "in a sense", "in a way", "in essence", "in one sense", "in particular", "kind of", "largely", "like", "likely", "literally", "look like", "looked upon", "looks like", "loosely speaking", "more or less", "mostly", "mutatis mutandis", "nominally", "occasionally", "often", "or", "par excellence", "particularly", "perhaps", "practically", "presumably", "pretty much", "principally", "quintessential", "rather", "really", "regarded", "relatively", "roughly", "seem", "seems", "so-called", "so to say", "sometimes", "somewhat", "sort of", "specifically", "strictly speaking", "such as", "technically", "tend", "tends", "typical", "typically", "usually", "very", "virtually"

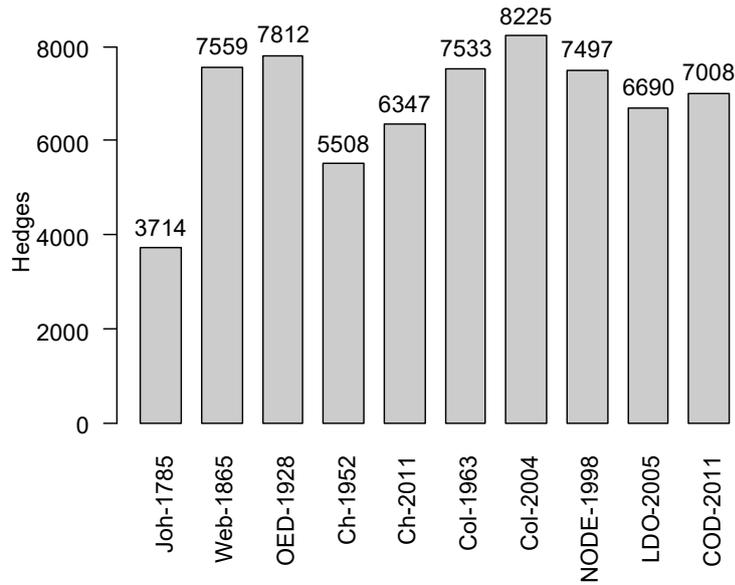
After identification of the hedges in each dictionary sample, their frequencies were computed and normalized per 100,000 with the aid of a specially designed computer program.<sup>10</sup>

### 2.3 Data analysis

The hedges were analyzed with respect to the overall frequency of the tokens and the distribution of the types across the dictionaries. The results of the former analysis are presented in Figure 1. As for the hedge types, their distribution is presented in Table 2. For easier identification of the correlation between the dictionaries and the hedges, the data were visualized using a correspondence analysis (CA) plot. CA produces a pictorial representation of the variables (here dictionaries and the hedges) on the same set of axes (Figure 2). The plot reveals associations between different data points on the basis of their co-occurrence (Glynn 2014: 133). The association is greater for the points located in proximity to one another, and away from the center of the plot. CA was performed on the hedges with the highest overall frequency of occurrence; to that end, the hedges ranked 1–30 were taken into consideration (see Table 2).

The CA plot indicates the most distinctive associations between the hedges and the dictionaries, but without showing whether the associations are statistically significant. For a more fine-grained view of these relations, a one-sided binomial test was performed on the same data, that is the first 30 rows of Table 2. The goal of the test was to verify the hypothesis that the probability of obtaining the observed value of hedge frequency in each cell is significantly higher than chance probability.<sup>11</sup> The resulting p-values for this hypothesis are displayed in Table 3. The values lower than the significance level 0.05, which are highlighted in grey, indicate that the frequencies are statistically higher than those expected by chance.

**Figure 1:** Overall frequencies of hedge tokens in the samples (normalized per 100,000)



**Table 1:** The size of the samples and percentage of hedges

	Johnson-1785	Webster-1865	OED-1928	Chambers-1952	Chambers-2011	Collegiate-1963	Collegiate-2004	NODE-1998	LDOCE-2005	COD-2011
Sample size (in word tokens)	16276	32685	16134	23550	28153	37664	35019	29444	16737	20530
Percentage of hedges	3.7	7.5	7.8	5.5	6.3	7.5	8.2	7.5	6.6	7.0

**Table 2:** Normalized frequencies of hedge types in the samples

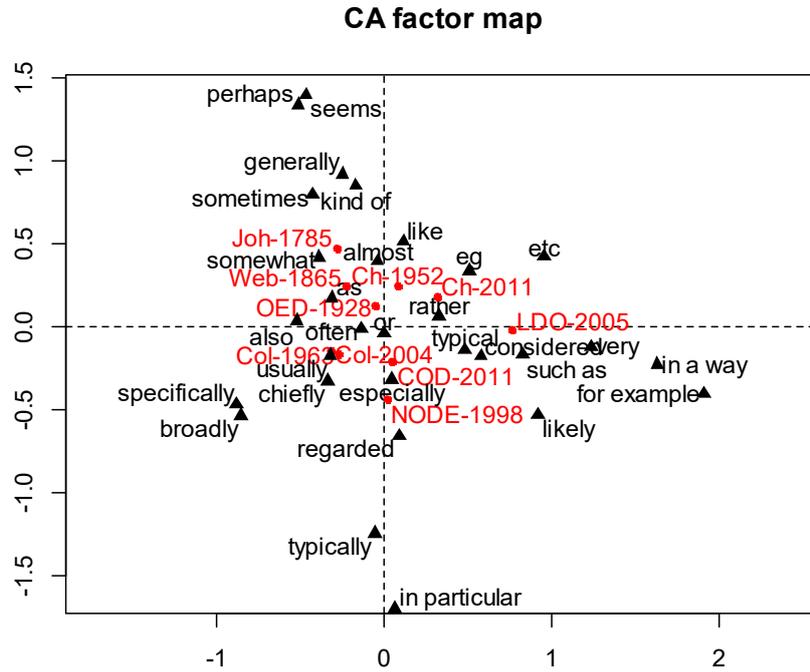
		Joh-1785	Web-1865	OED-1928	Ch-1952	Ch-2011	Col-1963	Col-2004	NODE-1998	LDO-2005	COD-2011
1	or	2255	4497	5014	3210	3694	4649	4532	5149	3866	4944
2	as	725	1787	862	769	586	1062	1471	672	341	653
3	especially	6	132	310	310	469	605	791	567	418	438
4	etc	0	128	359	327	668	0	0	31	633	117
5	also	49	162	359	59	60	244	371	48	0	49

6	like	55	285	217	268	170	45	23	54	102	97
7	usually	6	76	130	72	89	287	411	34	102	15
8	very	18	37	25	81	60	45	49	95	544	83
9	often	43	58	136	42	67	114	183	54	78	44
10	typically	6	0	0	4	7	27	46	282	0	156
11	sometimes	129	70	56	68	25	24	51	20	6	0
12	kind of	80	104	62	64	36	5	3	17	18	10
13	such as	37	3	12	8	60	3	0	71	114	73
14	eg	0	0	68	38	167	0	0	20	0	58
15	chiefly	6	31	74	13	4	48	51	54	0	63
16	specifically	0	21	0	0	4	138	77	0	0	0
17	generally	55	18	37	42	14	11	9	3	6	0
18	rather	37	3	0	8	21	16	17	24	42	24
19	considered	0	18	12	13	18	8	11	34	42	5
20	regarded	0	12	6	0	11	8	6	34	6	44
21	in particular	0	0	0	0	0	0	0	109	0	10
22	likely	0	0	0	0	4	21	6	17	42	0
23	for example	0	0	0	0	0	0	0	14	66	5
24	in a way	0	6	0	0	0	3	3	7	60	5
25	somewhat	6	15	25	8	4	19	0	3	0	0
26	seems	55	3	0	0	4	3	0	0	6	5
27	typical	0	0	6	13	4	8	9	3	18	15
28	perhaps	43	6	0	0	14	8	0	0	0	0
29	almost	18	6	0	4	7	8	0	3	6	15
30	broadly	0	0	0	4	0	42	20	0	0	0
31	really	0	0	0	8	4	0	0	0	48	0
32	particularly	18	15	0	0	4	0	0	17	0	5
33	relatively	0	6	0	0	0	5	26	10	0	10
34	actually	0	3	12	0	4	3	6	10	12	5
35	believed	0	0	0	13	0	5	0	3	18	15
36	more or less	6	3	12	13	0	5	3	7	0	0
37	seem	0	0	0	0	0	3	0	3	42	0
38	assumed	0	12	6	4	0	8	14	3	0	0
39	sort of	31	12	0	0	4	0	0	0	0	0



74	looked upon	0	0	0	0	0	0	0	0	0	0
75	loosely speaking	0	0	0	0	0	0	0	0	0	0
76	mutatis mutandis	0	0	0	0	0	0	0	0	0	0
77	nominally	0	0	0	0	0	0	0	0	0	0
78	par excellence	0	0	0	0	0	0	0	0	0	0
79	presumably	0	0	0	0	0	0	0	0	0	0
80	pretty much	0	0	0	0	0	0	0	0	0	0
81	quintessential	0	0	0	0	0	0	0	0	0	0
82	so to say	0	0	0	0	0	0	0	0	0	0
83	strictly speaking	0	0	0	0	0	0	0	0	0	0

**Figure 2:** Correspondence Analysis: Associations between dictionaries and hedges



**Table 3:** p-values for the binomial test calculated on the 30 most frequent hedges<sup>12</sup>

		Joh-1785	Web-1828	OED-1928	Ch-1952	Ch-2011	Col-1963	Col-2004	NODE-1998	LDO-2005	COD-2011
1	or	0.58	1	0.01	1	1	0.5	1	0	1	0
2	as	0	0	1	0.04	1	0.01	0	1	1	1
3	especially	1	1	1	0.85	0	0	0	0	0.1	0.16
4	etc	1	1	0	0	0	1	1	1	0	1
5	also	1	0.33	0	1	1	0	0	1	1	1
6	like	0.98	0	0	0	0	1	1	1	0.99	1
7	usually	1	1	0.86	1	1	0	0	1	0.96	1
8	very	1	1	1	0.65	1	1	1	0.98	0	1
9	often	0.6	1	0	1	0.89	0.01	0	1	0.59	1
10	typically	1	1	1	1	1	1	1	0	1	0
11	sometimes	0	0	0.3	0	1	1	0.72	1	1	1
12	kind of	0	0	0.01	0	0.61	1	1	1	1	1
13	such as	0	1	1	1	0	1	1	0	0	0
14	eg	1	1	0	0.04	0	1	1	1	1	0
15	chiefly	1	0.91	0	1	1	0.06	0.08	0.01	1	0
16	specifically	1	0.9	1	1	1	0	0	1	1	1
17	generally	0	0.83	0	0	0.88	1	1	1	1	1
18	rather	0	1	1	0.99	0.26	0.91	0.94	0.3	0	0.19
19	considered	1	0.53	0.97	0.55	0.25	1	0.99	0	0	1
20	regarded	1	0.77	1	1	0.65	0.98	1	0	0.99	0
21	in particular	1	1	1	1	1	1	1	0	1	0.8
22	likely	1	1	1	1	0.97	0	0.97	0.02	0	1
23	for example	1	1	1	1	1	1	1	0.08	0	0.95
24	in a way	1	0.92	1	1	1	1	1	0.83	0	0.94
25	somewhat	0.27	0.03	0	0.32	0.95	0	1	0.99	1	1
26	seems	0	0.99	1	1	0.93	0.99	1	1	0.76	0.9
27	typical	1	1	0.89	0.01	0.93	0.62	0.58	0.99	0	0.01
28	perhaps	0	0.82	1	1	0.01	0.54	1	1	1	1
29	almost	0	0.77	1	0.8	0.44	0.47	1	0.98	0.64	0
30	broadly	1	1	1	0.79	1	0	0	1	1	1

### 3. Results and discussion

The first point to make is that the hedges occur in the definitions in a relatively high proportion, which ranges from 3.7% in Johnson-1785 to 8.2% in Collegiate-2004 (see Table 1). As shown in Figure 1, they are used most frequently in the Collegiate (both editions), the OED, and Webster, and least frequently in Johnson and Chambers. The second point is that they occur statistically more frequently in the recent edition of the Collegiate than in the one from 1963 (8225 word tokens in Collegiate-2004 vs. 7533 in Collegiate-1963).<sup>13</sup> A similar finding emerges from the comparison of the two editions of Chambers (6347 word tokens in Chambers-2011 and 5508 in Chambers-1952).<sup>14</sup>

Analyzing the data in Tables 2 and 3, one finds that there is a great deal of variation in the use of hedges across the dictionaries. Although they are distributed rather unevenly in the dictionaries, the texts display preferences for particular sets of hedges. Johnson relatively rarely used hedges (3.7%), but it does not mean that he dispensed with them altogether. He arguably preferred *sometimes*, *kind of*, *sort of*, *such as*, *generally*, *rather*, *seems/seem*, *perhaps*, *almost*. They are characteristic of Johnson's defining style, as they occur significantly more often than one would expect by chance (see Table 3). Interestingly, Johnson appears to have a tendency to use hedges that point to less typical examples of a category (*sometimes*) or those extending the category boundary (*rather*, *almost*). For example, *sometimes* introduces secondary senses and less typical grammatical and semantic properties of words:

- farewell** The parting complement; adieu ... It is sometimes used only as an expression of separation without kindness.
- hight** ... 4. It is sometimes used as a participle passive: called; named.
- none** ... 4. *None of* sometimes signifies only emphatically *nothing*.
- virtue** ... 5. Efficacy ; power. Before *virtue* is used sometimes *by* and sometimes *in*

Looking at Table 2, one notices that *especially*, *typically* / *typical*, *usually*, which have become a standard feature of the metalanguage of many contemporary dictionaries, are rarely used in Johnson. According to Hanks (2005, 265), the lack of "and the like", "any of various", and "etc." should be attributed to Johnson's ability to "write pointed, no-nonsense definitions, seizing again and again on the central point."

In quantitative terms, Webster stands in contrast to Johnson, as the former has twice as many hedges as the latter (7559 and 3714, respectively) (Figure 1). The reasons lie in Webster's earlier edition of 1828, which had borrowed heavily from Johnson, using the latter as a "working base" for constructing the definitions (Friend 1967, 44-5). In this edition, a number of definitions were copied from Johnson verbatim or revised by adding more details on the basis of Johnson's examples. Webster's attention to details manifests itself by a more fine-grained division of senses<sup>15</sup> and provision of typical contextual uses of a word

being defined. Roughly speaking, what Johnson provides in illustrative quotations Webster (1828) includes in definitions following the hedge *as*:

Johnson-1785: **nice** ... 6. Requiring scrupulous exactness.

*Supposing an injury done, it is a nice point to proportion the reparation ...*

L'Estrange

Webster-1828: **nice** ... 4. Requiring scrupulous exactness; as a *nice* point.

The 1864 edition of Webster, which is under study<sup>16</sup>, was revised substantially, but the use of *as* introducing typical examples survived:

Webster (1865): **nice** ... 3. ... to be scrupulously and exactly handled; delicate; refined; as, a *nice* distinction of point.

Worth noting is that *as* is distinctively associated with both Webster and its recent descendant, the Collegiate (see Table 2).

The dictionary that comes second in the ranking (Figure 1) is the OED. One of the reasons for this dictionary being rich in hedges is certainly the fact that it is a historical dictionary. The methodology and the purpose of compiling historical dictionaries make hedges inevitable. Historical lexicographers examine corpus evidence and peruse countless uses of a lexical unit in order to determine one or more prototypical core senses and then possible lines of development of secondary and peripheral senses (see also Kay 2000). The aim is to give a comprehensive account of the history of lexical units by uncovering the full path of their semantic change. A fully thorough treatment of meaning is impossible to accomplish when the description is to be rendered in terms of necessary and sufficient conditions. For one thing, indeterminacy of meaning causes demarcation problems, which can only be solved by focusing on what is typical in the semantic structure. It is more convenient to use hedges than to attempt to enumerate the entire range of category members or features. This strategy can be seen in the definition of **head** below, which mentions *ale* and *beer* as typical examples of liquor that bears a *head*, while omitting *palm-wine* which is nevertheless recorded in one of the citations (see below).

**head** 10. A collection of foam or froth on the top of liquor, esp. ale or beer.

... 1760–72 ... *Palm-wine .. bears a greater head than beer, and is of a very inebriating quality* (OED)

The frequent use of hedges is a distinctive characteristic of the definition style, which is especially noticeable in the dictionaries that aim to serve as a historical record of the language. They aim to describe the entire semantic structure of lexical units by providing dominant and peripheral senses, along with less frequent and more specialized readings. Hedges along with an elaborate sense division (e.g. I, II, III, 1, 2, 3, a, b, c), albeit a linear one, are lexicographic means of capturing the prototype (Geeraerts 2006, 357; Zgusta 2006, 114).

While Johnson uses relatively few hedges, Collegiate-2004 resorts to them

most often. The fact that it has the largest total number of hedges is arguably due to the developed system of sense division, with hedges being used as sense dividers. Along with Arabic numerals, letters, parenthesized numerals, and colons, which indicate various levels of the hierarchy of senses, the dictionary uses *esp(ecially)*, *specif(ically)*, *also*, and *broadly* to indicate "a particular semantic relationship" between (sub-)senses (The Collegiate 2003, 20a). In addition, it employs a hedge *as* to indicate "that the following subsenses are typical or significant examples" (*ibid.*), for example:

- billow** ... 1 : WAVE; *esp* : a great wave or surge of water 2 : a rolling mass (as of flame or smoke) that resembles a high wave
- billy club** ... a heavy usu. wooden club; *specif* : a police officer's club
- binary** ... something made of or based on two things or parts: as **a** : binary star **b** : a binary number system
- outgas** 1 : to remove occluded gases from usu. by heating; *broadly* : to remove gases from
- putting green** a smooth grassy area at the end of a golf fairway containing the hole; *also* : a similar area usu. with many holes that is used for practice

As Table 3 demonstrates, the above hedges occur in both editions of the Collegiate significantly more often than chance would predict. They are distinctively associated with this dictionary. In addition to the hedges indicating relations between particular senses, the Collegiate (in both editions) displays a significant preference for *usu(ally)* and *often* as indicators of typicality within definitions (e.g. **aspersion** "a sprinkling with water *esp.* in religious ceremonies" Collegiate-2004). Neither of the above uses is original in lexicography, as they can be traced back to earlier dictionaries, notably Webster, of which the Collegiate is a remote descendant.<sup>17</sup> Nevertheless, as far as sense-division is concerned, the Collegiate has exploited hedges to a remarkable extent.

Of all the popular dictionaries, Chambers uses the least number of hedges. One of the reasons for this is a marked tendency to use brief definitions, a strategy resulting from the policy of all-inclusiveness. Brief definitions, often in the form of a string of near-synonyms, which are a distinctive feature of Chambers, leave little room for prototypical information. The comprehensive coverage of meanings, which Chambers aims at, stems from the desire to make the dictionary useful in reading a vast range of texts. Chambers' aim of comprehensiveness is pursued by the coverage of literary words and senses used by Shakespeare, Spenser, and Milton, as well as the vocabulary pertaining to non-English varieties of English (e.g. Scottish English). The need to provide the user with peripheral, rare, and less predictable meanings outweighs the need to give the details of prototypical meanings, which the native user is familiar with (cf. Geeraerts 2006: 330). Yet the two editions of Chambers have their own distinctive set of hedges: *etc*, *like*, and *e.g.* occur significantly more frequently in this dictionary than one may expect by chance (see Table 3). Unlike the Collegiate, Chambers

does not use hedges as markers of the hierarchy of senses.<sup>18</sup>

The other popular dictionaries, COD-2011 and NODE-1998, share a few characteristic hedges: *or, typically, such as, chiefly, regarded* (see Table 3). These are arguably traces of the same methodology of preparing definitions; note that both dictionaries were compiled by the same editor and that the former dictionary drew heavily on the latter (Preface to COD-2011). However, NODE-1998 has a few other hedges that are distinctively associated with this dictionary: *especially, considered, likely, and in particular*. The last expression is noteworthy, as it is used systematically to introduce specialized cases of a core sense, as in:

- crossbar** ... a horizontal bar fixed across another bar or between two upright bars, in particular:
- the bar between the two upright posts of a football goal.
  - the horizontal metal bar between the handlebars and saddle on a man's or boy's bicycle.

It is the only hedge used in NODE-1998 consistently to specify the type of relationship that holds between a core sense and one or more sub-senses.

As the only dictionary for foreign learners, LDOCE-2005 has its own distinctive set of hedges, which occupy the most extreme positions in the right section of Figure 2, relatively close to this dictionary. It is clear that the dictionary displays a marked preference for *etc., very, such as, rather, considered, likely, for example, in a way, typical, seem, definitely* (see also Tables 3 and 2). These words and expressions belong to the restricted defining vocabulary, which is intended to make the definitions accessible to the learner. They were selected largely according to the criterion of frequency and range in the language, in order to ensure that their form and meaning is straightforward and leaves little room for misinterpretation. This policy differs radically from that employed in some native-speaker dictionaries which permit words in peripheral meanings, for instance, while LDOCE frequently uses *for example* to introduce prototypical examples, Johnson, Webster, and the Collegiate consistently use *as* to this end.

Finally, it is worth noting that 25% of the hedges (i.e. 21 hedge types) under consideration are not used in any definition (see Table 2). This can be explained by the fact that some of them, such as *mutatis mutandis, par excellence, quintessential, in a manner of speaking* are too formal to be admitted into definitions in a systematic way. In turn, *pretty much* and *so to say* are excluded due to their informality.

#### 4. Conclusions

The foregoing analysis suggests several conclusions. Firstly, the fact that hedges are used more extensively in the recent editions of the popular dictionaries (i.e. Chambers-2011 and the Collegiate-2004) than in the editions published prior to the 1970s suggests that contemporary defining practices in these

dictionaries must have been inspired by prototype theory. Being part of the defining metalanguage, they are imposed on lexicographers by editors through dictionary style manuals. Their use is more conscious and controlled now than in the past, with the aid of dictionary writing systems, which help the editor manage the production of a dictionary and execute consistently the dictionary policy (Abel 2012). Besides being used as indicators of prototypical elements of meaning within a definition, hedges are applied to highlight sense relationships. While the former use has been widespread in the English dictionaries, the latter is not.

Prototype theory, however, is not the only factor determining the use of hedges. The analysis of the sample shows that while the intensification of the use of hedges in the recent popular dictionaries should be attributed to the influence of this theory on dictionary editors, the high rate of hedges in early dictionaries (notably the OED-1928 and Webster-1865) suggests that other factors must be at play. A few aspects relevant to dictionary function (Tarp 2008) emerge as key factors: the depth of treatment, dictionary type, and defining policy. Hedges are particularly useful in large-scale dictionaries compiled on historical principles (such as the OED), as the full account of the history of word meanings would be impossible to achieve without them. They are also useful in desk-size dictionaries that aim at thoroughness of treatment (cf. Geeraerts 1997). Their frequent use in Webster-1865 results partly from the systematic strategy of combining Johnson's definitions and examples by means of "as". On the other hand, they are less common in dictionaries that rely heavily on brief and synonymous definitions (such as Chambers). Other factors such as individual lexicographers' preferences and the stylistic guidelines as specified in dictionary style manuals cause that different dictionaries show marked preferences for different sets of hedges.

Finally, it is hardly surprising that hedges have been used more or less extensively in the English dictionaries since Johnson-1785. Arguably, the well-established use of hedges in English lexicography results from their usefulness, if not necessity: they enable the definer to achieve what otherwise would be difficult to achieve: economy of expression, as well as reliability and thoroughness of description.

## Endnotes

1. John Taylor mentions examples of adjectives (*true, real*), suffixes (*-ish*), conjunctions (*in that, and, or*), certain lexical and auxiliary verbs (*seem, appear, may, can*), and even inverted commas (1995: 76).
2. However, the overuse of hedges in definitions has been criticized by Anna Wierzbicka, who considers them as "visible signs of indecision and analytical failure" (1996: 269).
3. The former specifies the class to which an object being defined belongs, and the latter, the features distinguishing the object from other members of the class (Svensén 1993).

4. While the classical definition is effective in defining large sections of the lexicon, it is clearly not suitable for those areas of lexis which display taxonomic gaps (Adamska-Salaciak 2012).
5. A much-quoted definition of **door** in *Webster's Third International* (1961) occupies 11 lines of text (Atkins and Rundell 2008).
6. The sampling was aided with an R function called *sample*, which allowed for a random selection of page numbers. Initially, 30 pages from each dictionary were drawn. However, due to the different density of the dictionaries texts, further pages were selected from the dictionaries with the lower text density in order to reduce differences in sample sizes. The error caused by the different sample sizes was further minimized by the normalization of hedge frequencies.
7. The images were processed with ABBYY FineReader 11.
8. Lakoff's list (1973: 472) has 67 items collected under the heading "Some hedges and related phenomena".
9. Among the hedges excluded from the study were those that were computationally intractable (e.g. "he's another (Caruso/Lincoln/ Babe Ruth/ ...)") and those likely to produce false hits (e.g. a suffix "-ish", which is often used in the nationality sense, as in *English*).
10. The program was written in R, an open source programming language (R Development Core Team 2013).
11. The binomial test was conducted with the *binom.test* function in R, with the *alternative* argument set to "greater" for a one-sided test. The test calculated for each cell the probability of the observed value given the expected probability. The latter was computed by multiplying the column frequency total by the row frequency total and dividing the result by the total frequencies in the table (for the rows 1-30) (see also Gries and Stefanowitsch 2006; Janda 2013; Stefanowitsch and Gries 2003).
12. The values equal 0 or lower than the significance level 0.05 indicate that the frequencies are statistically higher than those expected by chance.
13. The difference is statistically significant, as a chi-square test gives  $p < 0.05$ ,  $\chi^2 = 30.389$ ,  $df = 1$ .
14. A chi-square test gives  $p < 0.05$ ,  $\chi^2 = 59.378$ ,  $df = 1$ .
15. For example, while Johnson provides one sense at **hospitable**, Webster distinguishes three. At **nice**, there are 9 and 14 senses, respectively.
16. Although the current study was conducted on the 1865 version of Webster, it is the same edition as that published in 1864.
17. The Collegiate series derives from the unabridged *Webster's International Dictionary* (1890) (and its revisions), which in turn is a direct descendant of Webster's famous revision of 1864.
18. In Chambers-1952 sense arrangement generally adheres to the historical principle, whereby the first definition describes the "original" or early meaning, and the following ones cover a series of secondary meanings "branching out and diverging from" the original one (Chambers-1952: vii). Although the linear presentation of senses, separated by colons, is a practical space-saving strategy, it creates a simplified picture of sense relations, implying that secondary meanings derive directly from one and the same etymological root. A more sophisticated presentation of senses, arguably inspired by prototype theory, is found in COD-2011. The dictionary organizes senses according to typicality, providing more than one core sense and subsenses deriving from the core.

## References

- Abel, A. 2012. Dictionary Writing Systems and Beyond. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography*: 83-106. Oxford: Oxford University Press.
- Adamska-Sałaciak, A. 2012. Dictionary Definitions: Problems and Solutions. *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 129(4): 7-23.
- Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.
- Friend, J.H. 1967. *The Development of American Lexicography 1798–1864*. The Hague: Mouton.
- Geddie, W. 1952. *Chambers's Twentieth Century Dictionary*. Second edition. Edinburgh: W. & R. Chambers.
- Geeraerts, D. 1997. *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. Oxford: Clarendon Press.
- Geeraerts, D. 2006. The Definitional Practice of Dictionaries and the Cognitive Semantic Conception of Polysemy. Geeraerts, D. (Ed.). 2006. *Words and other Wonders. Papers on Lexical and Semantic Topics*: 345-363. Berlin: Mouton de Gruyter.
- Glynn, D. 2014. Correspondence Analysis: Exploring Data and Identifying Patterns. Glynn, D. and J. Robinson (Eds.). 2014. *Corpus Methods for Semantics. Quantitative Studies in Polysemy and Synonymy*: 443-486. Amsterdam: John Benjamins.
- Gove, P. (Ed.). 1963. *Webster's Seventh New Collegiate Dictionary*. Seventh edition. Springfield, Mass.: G. & C. Merriam. (Collegiate-1963)
- Gries, S.Th. and A. Stefanowitsch (Eds.). 2006. *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. Berlin: Walter de Gruyter.
- Hanks, P. 1994. Lexicography: Theory and Practice. *Dictionaries* 14(1992/3): 97-112.
- Hanks, P. 2005. Johnson and Modern Lexicography. *International Journal of Lexicography* 18(2): 243-266. <https://doi.org/10.1093/ijl/eci024>.
- Hyland, K. 1994. Hedging in Academic Writing and EAP Textbooks. *English for Specific Purposes* 13(3): 239-256. [https://doi.org/10.1016/0889-4906\(94\)90004-3](https://doi.org/10.1016/0889-4906(94)90004-3).
- Hyland, K. 2006. Medical Discourse: Hedges. Brown, K. (Ed.). 2006. *Encyclopedia of Language and Linguistics*: 694-697. Second edition. Oxford: Elsevier.
- Janda, L.A. 2013. Quantitative Methods in Cognitive Linguistics: An Introduction. Janda, L.A. (Ed.). 2013. *Cognitive Linguistics: The Quantitative Turn*. Berlin: De Gruyter Mouton.
- Johnson, S. 1785. *A Dictionary of the English Language*. [Vol. 1 and 2]. London: J.F and C. Rivington. (Johnson-1785).
- Kay, J.C. 2000. Historical Semantics and Historical Lexicography: Will the Twain Ever Meet? Coleman, J. and C.J. Kay (Eds.). 2000. *Lexicology, Semantics and Lexicography*: 53-68. Amsterdam: John Benjamins.
- Lakoff, G. 1973. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic* 2(4): 485-508.
- Lakoff, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Landau, S.I. 2001. *Dictionaries: the Art and Craft of Lexicography*. Second edition. New York/ Cambridge: Cambridge University Press.
- Landau, S.I. 2009. Major American Dictionaries. Cowie, A.P. (Ed.). 2009. *The Oxford History of English Lexicography. Vol. 1*: 182-229. Oxford: Oxford University Press.

- Langacker, R.W.** 1987. *Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Lewandowska-Tomaszczyk, B.** 2007. Polysemy, Prototypes, and Radial Categories. Geeraerts, D. and H. Cuyckens (Eds.). 2007. *The Oxford Handbook of Cognitive Linguistics*: 139-169. Oxford: Oxford University Press.
- Marr, V. (Ed.)**. 2011. *The Chambers Dictionary*. Twelfth edition. London: Chambers Harrap. (Chambers-2011)
- Mish, F. (Ed.)**. 2004. *Merriam-Webster's Collegiate Dictionary*. Eleventh edition. Springfield, Mass.: Merriam-Webster. (Collegiate-2004)
- Murray, J.A.H. (Ed.)**. 1888–1928. *The Oxford English Dictionary on Historical Principles*. Vol. 1–10. [Edited by James Murray, Henry Bradley, William A. Craigie, and Charles. T. Onions]. Oxford: Clarendon Press. (OED-1928)
- Pearsall, J. (Ed.)**. 1998. *The New Oxford Dictionary of English*. Oxford: Clarendon Press. (NODE-1998)
- R Development Core Team**. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rey, A.** 1990. Definitional Semantics: Its Evolution in French Lexicography. Tomaszczyk, J. and B. Lewandowska-Tomaszczyk (Eds.). 1990. *Meaning and Lexicography*: 43-55. Amsterdam/Philadelphia: John Benjamins.
- Rosch, E.** 1973. Natural Categories. *Cognitive Psychology* 4: 328-350.
- Rosch, E.** 1975. Universals and Cultural Specifics in Human Categorisation. Brislin, R.W., S. Bochner and W.J. Lonner (Eds.). 1975. *Cross-cultural Perspectives on Learning*: 177-206. New York: Halstead/Wiley.
- Stefanowitsch, A. and S.Th. Gries**. 2003. Collostructions: Investigating the Interaction of Words and Constructions. *International Journal of Corpus Linguistics* 8(2): 209-243.
- Stevenson, A. and M. Waite (Eds.)**. 2011. *Concise Oxford English Dictionary*. Twelfth edition. Oxford: Oxford University Press. (COD-2011)
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford/New York: Oxford University Press.
- Swanepoel, P.** 1994. Problems, Theories and Methodologies in Current Lexicographic Semantic Research. Martin, W. et al. (Eds.). 1994. *Euralex 1994. Proceedings* (Beiheft): 11-26. Amsterdam: Vrije Universiteit.
- Summers, D. et al. (Eds.)**. 2005. *Longman Dictionary of Contemporary English*. Fourth Edition with Writing Assistant. Harlow: Pearson Education. (LDOCE-2005)
- Tarp, S.** 2008. *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Max Niemeyer.
- Taylor, J.R.** 1995. *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.
- Van der Meer, G.** 1999. Metaphors and Dictionaries: The Morass of Meaning, or How to Get Two Ideas for One. *International Journal of Lexicography* 12(3): 195-208. <https://doi.org/10.1093/ijl/12.3.195>.
- Van der Meer, G.** 2000. Core, Subsense, and the *New Oxford Dictionary of English* (NODE). On How Meanings Hang Together, and Not Separately. Heid, U., S. Evert, E. Lehmann and C. Rohrer (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, August 8th–12th, 2000*: 419-431. Stuttgart: Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.

**Webster, N.** 1828. *An American Dictionary of the English Language*. New York: S. Converse. (Webster-1828)

**Webster, N.** 1865 [1864]. *An American Dictionary of the English Language*. [Revised by C.A. Goodrich and N. Porter.] Springfield, Mass.: Merriam. (Webster-1865)

**Wierzbicka, A.** 1996. *Semantics. Primes and Universals*. Oxford: Oxford University Press.

**Zgusta, L.** 2006. *Lexicography Then and Now. Selected Essays*. [Edited by Fredric S.F. Dolezal and Thomas B.I. Creamer]. Tübingen: Max Niemeyer.