

# Neoclassical compounds and final combining forms in English

Ana Díaz-Negrillo (Granada)

---

## Abstract

English neoclassical compounds rely on a distinct vocabulary stock and present morphological features which raise a number of theoretical questions. Generalisations about neoclassical compounds are also problematic because the output is by no means homogeneous, that is, defining features of neoclassical compounds sometimes co-exist with features that are not prototypical of these formations. The paper looks at neoclassical compounds with a view to exploring patterns of morphological behaviour and development in this class of compounds. The approach is both synchronic and diachronic: it researches whether the morphological behaviour of recently formed compounds is different from that of earlier compounds and, if so, in which respects. This is assessed on data from the BNC with respect to some of the features that are cited in the literature as defining properties of neoclassical compounds, specifically, their internal configuration, the occurrence or not of a linking vowel, and their productivity.

---

## 1 Introduction

Neoclassical formations are noted for using a large part of the Greek and Latin vocabulary stock that exists in English. In particular, they use the so-called combining forms (hereafter CFs), which are bound morphemes that in principle differ from bound roots and affixes, even if this distinction is difficult to draw in many cases. In addition to CFs, neoclassical compounds often have a linking vowel in medial position between the bases of the compounds. A final, crucial feature of neoclassical compounds is that they are an active source of vocabulary extension, that is, they are productive nowadays.

Despite these properties, the words listed in the literature as neoclassical compounds are by no means uniform. Neoclassical compounds may exhibit a variety of configurations, first in terms of the origin and morphological status of their constituent elements and, second, in terms of the occurrence of a linking vowel or not. Finally, not all types of neoclassical compounds, and their internal configurations, appear to be equally productive.

This paper aims at assessing the morphological behaviour and the development of neoclassical compounds with respect to the above defining properties of neoclassical compounds: i) the combinatorial possibilities of their constituent elements, ii) the occurrence or not of a linking vowel, and iii) their productivity. A quantitative exploration of the incidence of those properties may cast light on the morphological behaviour of the words that are usually described as neoclassical compounds. This paper uses synchronic analysis for any current tendencies in their behaviour, and diachronic analysis for any evidence of the various ways in which the formations have developed and, if available, for hints at morphological tendencies in this type of compounds. In the latter case, the aim is to find out whether the morphological nature of recently formed compounds is different from that of earlier compounds and, if so, in which respects.

For the synchronic analysis, the paper relies on data from 425 neoclassical compounds extracted from the *British National Corpus* (BNC) classified according to 10 final combining forms (hereafter FCFs), that is CFs that stand in final position in the compound. For the diachronic

analysis, the paper uses the earliest attestations of the compounds under study according to the *Oxford English Dictionary* (hereafter OED).

The paper is structured into this Introduction and another six sections. Section 2 reviews the defining properties of neoclassical compounds and of CFs. Section 3 describes the method and the data. Sections 4 to 6 go into the three points under study: the analysis of the combinatorial possibilities of FCFs (section 4), the occurrence of a linking vowel (section 5), and the productivity of the neoclassical compounds in the study by FCFs (section 6). Each of these sections provides a synchronic and diachronic gradience of the FCFs in the study with respect to each of these three points. Section 7 discusses the results and section 8 summarizes the conclusions of the study.

## 2 Neoclassical compounds and combining forms

Neoclassical compounds are formations that consist of at least one CF. CFs were lexemes in the classical languages. Their lexemic status can be seen, for example, in that semantic correspondences between bound roots and free native morphemes can be established, e.g. *pedo-* 'child', *-lith* 'stone', *-ectomy* 'excision'. In terms of autonomy, CFs are bound, i.e. they cannot stand as free lexemes and have no free morphologically-related correspondents in English. Accordingly, they have also been called *stems* or *roots*, both terms used with similar senses. Dieter Kastovsky (2009: 9–10) argues that the class *stem*, which contains elements like *scient-* (as in *scient-ist*), covers the constituent elements occurring in neoclassical compounds (for Dieter Kastovsky, the other two types of inputs of English morphological processes are words and clipped forms). Valerie Adams (2001: 110) also refers to CFs as *stems* meaning 'bound lexical bases', and she uses *stem compounds* to refer to what we call here *neoclassical compounds*.<sup>1</sup> Geert Booij (1992: 56) refers to CFs as *roots* and calls *non-native compounds* the type of formations under study here.<sup>2</sup>

By contrast, Laurie Bauer (1983: 213–216) justifies the existence of the class *combining form* on the grounds of the combinatorial possibilities of its members. Although bound roots and CFs share their boundness, roots can combine with suffixes to form free words, but CFs do not form words if combined with suffixes (cf. also Warren 1990: 122). Among the combinatorial possibilities of CFs, the structure [ICF + FCF], where ICF stands for *initial combining form*, is the central compound type, e.g. *astronaut*, *fratricide* (cf. Bauer/Huddleston 2002: 1661).<sup>3</sup> Still, as remarked in the literature, CFs in neoclassical compounds can also combine with bound roots (cf. Plag 2003: 156), e.g. *glaci-* in *glaciology*, with free roots (cf. Bauer/Huddleston 2002: 1662), e.g. *merit* and *electric* in *meritocrat* and *hydro-electric*, and with clipped words, e.g. *Euro* in *Eurocrat* (cf. Bauer 1998: 408).<sup>4</sup> To our knowledge, the extent to which the prototypical configuration is commoner than other configurations has not been quantitatively explored in the literature.<sup>5</sup>

<sup>1</sup> The term *stem* is also used in the literature to refer to bound elements with a syntactic category membership (cf. Giegerich 1999: 88). In Heinz Giegerich's model of stratified grammar, syntactic category membership draws the difference between stems and bound roots.

<sup>2</sup> For a discussion of different views on the status of the constituent units of neoclassical compounds see, for example, Anke Lüdeling/Tanja Schmid/Sawwas Kiokpasoglou (2002: 257–258) or Dany Amiot/Georgette Dal (2007: 324–326).

<sup>3</sup> Some CFs can stand both in initial and final position, e.g. *lith* in *lithograph* or *megalith*, as opposed to *bio-* or *-ectomy*, which only take initial and final position, respectively.

<sup>4</sup> Bauer (1998: 408) explains that *Eurocrat* may have more than one analysis. In addition to seeing it as a clipping of *European* added to the CF *-crat*, it could be analysed alternatively as a clipping added to a splinter from *bureaucrat*, or as a blend from *European* and *bureaucrat*.

<sup>5</sup> As suggested by one of the reviewers, various theories of confixes influence the distinction between bound roots and clipped words (cf. for example, Kirkness 1995; Iacobini 2000; Stanforth 2005; Donalies 2009).

Another major feature of CFs is their classical origin. As mentioned earlier, CFs were lexemes in either Greek or Latin, where they were inflected and were also used for derivation and compounding. Their classical origin and their boundness become particularly relevant features for the implications that native vs. non-native, and free vs. bound have in English morphology. In a model of stratified grammar (cf. Siegel 1974; Allen 1978; Giegerich 1999), *bound* and *non-native* are features associated with a component of the lexicon which is governed by its own system of word-formation rules (cf. Kiparsky 1982; Aronoff/Fuhrhop 2002). Accordingly, and prototypically, classical CFs combine with other classical CFs. However, classical CFs often combine with free native bases too, e.g. *ufology*. As a result, Laurie Bauer (1998) refers to neoclassical compounds as a compromise type in a gradient model of English word-formation which develops along three axes: native vs. foreign; simplex-derivative-compound, and abbreviated vs. non-abbreviated. Nevertheless, the extent to which CFs combine with non-native bases in English neoclassical compounds has to our knowledge not been explored to date.

The occurrence of linking vowels in neoclassical compounds is also explained in the classical origin of CFs. The linking vowels that are frequent in neoclassical compounds go back to classical thematic vowels. They are often *-o-*, as in *epistemology*, and sometimes *-i-* as in *herbicide*. The analysis of these linking elements is not uncontroversial and is more relevant than it may appear. Its analysis often depends on whether the initial element is bound or free. If the compound contains a bound initial base, sometimes the element that stands as a linking element is actually part of the initial CF historically, as in *arachnophobia*. If the initial base is free, as in *rodenticide*, the claim that the middle vocalic element belongs to the initial base is questionable because, as Dieter Kastovsky (2009: 7) suggests, it may imply that an allomorph *rodenti-* exists (cf. however, Baeskow 2004: 99; Prčić 2008: 8). In cases like this latter, it seems more plausible to analyse the linking element as part of the last element (*-icide*, cf. Bauer/Huddleston 2002: 1663 specifically on *-icide*), or as part of neither element (cf. Kastovsky 2009: 6 on *-o-*logy). The occurrence of the linking vowel also seems to depend on specific CFs. It can be seen that some CFs preclude the occurrence of linking elements, especially the ICF or FCFs that end or start with vowels, respectively, e.g. *tele-*, *-ectomy*, whereas others seem to take one and the same linking vowel in the majority of cases, e.g. *(o)logy* or *(i)cide*. The choice of one or another linking vowel is also governed by the Greek (*-o-*) or Latin (*-i-*) origin of the FCFs in the compound. The literature cited above discusses the various possible analyses of the linking vowels in neoclassical compounds but, again, and to our knowledge, a quantitative analysis of the presence of the linking elements has not been undertaken to date.

Finally, neoclassical compounds are productive in present-day English (Bauer 1983: 216; Bauer/Huddleston 2002: 1661), which means that speakers are able to identify the constituents in these compounds and use them productively to form new neoclassical compounds. This is again interesting from the theoretical point of view, because productivity in the non-native component of the English language has often been questioned (cf. for example, Marle 1985). It also seems that some CFs enter neoclassical compounds more readily than others. Productivity becomes relevant in neoclassical compounds because of the effects that it can have in the morphology, and subsequent categorization, of the so-called CFs, especially if productivity is considered together with their combinatorial preferences, i.e. other bound or free bases. To the best of our knowledge, productivity in neoclassical compounds or across CFs has not been measured as yet, and therefore stands as the third research point of this paper.

Neoclassical compounds are often used in specialised registers. Thus, *-ectomy* is found mainly in Medicine terms and *-lith* in Biology and Pathology terms. As a result of their specialised use, it is often the case that outsiders of these disciplines have to look up neoclassical compounds, or their constituents, in terminological dictionaries. However, not all neoclassical compounds are as specialised and, therefore, infrequent in everyday language. Anke Lüdeling/Stefan Evert (2005), in relation to German *-itis*, which in Medicine means 'inflammation of a particular body

part', reported that it has recently become particularly productive with the non-specialised meaning 'excessive or in excess', e.g. *Telefonitis*. This means that productivity in neoclassical word-formation may actually happen in extended non-specialised uses. This is an area which may throw light on the evolution of CFs. Although this point is not within the scope of this paper, preliminary remarks will be made below based on the results obtained here.

Overall, the picture that emerges from the above is that neoclassical compounds form a heterogeneous class. Their heterogeneity manifests itself in the existence of output where CFs combine with other CFs but also with free roots or clipped words; or where linking elements occur in the compounds, or do not occur at all. In addition, some CFs are more likely to participate in new neoclassical compounds than others. Laurie Bauer (1998: 409) has argued that the class of neoclassical compounds is actually "a kind of prototype, from which actual formations may diverge in unpredictable ways". He goes on to argue that, although necessary as a class for the large number of elements it covers, neoclassical compounds should also be treated as part of a continuum, therefore having fuzzy borderlines with other categories. Similarly, Dany Amiot/Georgette Dal (2007), for French and following Claudio Iacobini (2004), for Italian, have claimed that, even though a number of central features may exist, each CF requires individual analysis. This paper aims at a quantitative analysis of individual FCFs, with the aim of disclosing as detailed evidence as possible of the morphological tendencies in the class.

### 3 Data collection and general figures

This paper draws on neoclassical compounds classified by their FCFs. The selection of a sample of FCFs is a thorny issue, considering the partial disagreement on the concept 'combining form', and also between lists of CFs published in the specialized literature. In order to minimize the bias, three of the main references that list CFs were used for the selection of FCFs, namely Beatrice Warren (1990), Ingo Plag (2003) and Laurie Bauer/Rodney Huddleston (2002: 1621–1721). The FCFs listed in at least two of these three references were selected. Suffixed FCFs were excluded.<sup>6</sup> After the application of these criteria 10 FCFs remained. They are listed in Table 1:

<i>cide</i>	<i>ectomy</i>	<i>logy</i> <sup>7</sup>	<i>morph</i>	<i>phobia</i>
<i>crat</i>	<i>lith</i>	<i>mania</i>	<i>phile</i>	<i>scope</i>

Table 1: Alphabetical list of the 10 FCFs in the study.<sup>8</sup>

The FCFs in the study sample illustrate the heterogeneity which is often associated with neoclassical compound and their building units. Three FCFs can be used as free-standing morphemes in contemporary English (*mania*, *phobia* and *scope*), and another three FCFs can be used also as ICFs (*lith*, *morph* and *phile*). While the morphological behaviour of the latter group will not be further discussed in the paper, the possible free status of the former group will become relevant in the interpretation of the morphological behaviour of FCFs.

<sup>6</sup> The initial selection included *-graphy* and *-cracy*. Arguably, *-y* is here an independent affix that is added to more basic and free complex units ending in *-graph* and *-crat*, as, for example, the stress-shift it imposes on the base suggests.

<sup>7</sup> Note, however, that *-logy* has not been disregarded because, as opposed to *-graphy* or *-cracy* and unlike French *-logue* or Spanish *-logo*, *-log* is not a possible FCF in English. Following Laurie Bauer/Rodney Huddleston (2002: 1665–1666), *-logy* formations can be considered the most basic form of derived compounds in the derivational paradigm (an anthropologist is a person who pursues the science of anthropology; anthropological is of, pertaining to, or connected with, anthropology, etc.).

<sup>8</sup> Some of the FCFs are cited in the source references with a vowel attached to them, in particular *-icide*, *-(o)logy* and *-ophile* (cf. Bauer/Huddleston 2002: 1661). All of them are cited here without the linking vowel. This is an aspect that will be empirically explored in the study (cf. sections 5 and 7.2).

Every two-base compound containing underived bases and ending in any of these FCFs was then retrieved from Adam Kilgarriff's unlemmatised list of the entire BNC (cf. Kilgarriff 1996). The BNC contains 100 millions words from texts in British English between the late 1980s and 1993. No distinctions were made for register or medium.

The online versions of the OED and the BNC (<http://corpus.byu.edu/bnc/>) were extensively used in order to disambiguate cases which may hold mere formal coincidence with the endings in the study, and also to verify the meaning of the compound. For example, *chatricide*, which is not listed in the OED, was disambiguated with the BNC: "[...] is the first recorded victim of chatricide. He has been chattered to death". The formations whose meaning could not be verified were discarded, e.g. *weidoscope*.

The final number of compounds collected for the study and their token frequencies are in Table 2:

	Types	%	Tokens	%
<i>cide</i>	31	7.29	2614	6.92
<i>crat</i>	10	2.35	1294	3.43
<i>ectomy</i>	39	9.18	579	1.53
<i>lith</i>	9	2.12	92	0.24
<i>logy</i>	177	41.65	31194	82.61
<i>mania</i>	48	11.29	122	0.32
<i>morph</i>	13	3.06	60	0.16
<i>phile</i>	25	5.88	127	0.34
<i>phobia</i>	29	6.82	334	0.88
<i>scope</i>	44	10.35	1346	3.56
	425		37762	

**Table 2: Distribution of the study sample by FCFs in types and tokens, with indication of their frequencies and percentages within the respective total number of types and tokens in the sample.**

For the study of the morphological development of the neoclassical compounds and FCFs of the study sample, the earliest record for every compound in the OED was collected. This information makes it possible to explore diachronically the morphological features under study. Although listedness in the OED may not necessarily coincide with coinage, the OED is probably one of the most reliable sources to pin down word coinage. The compounds that are not listed in the OED are treated as 20<sup>th</sup> century formations (47.6 % of the units in the 20<sup>th</sup> century), so the date information for these cases is that of the BNC corpus. Still, although this set of compounds is analysed among 20<sup>th</sup> century compounds, they will be presented separately from 20<sup>th</sup> century compounds which are listed in the OED. Table 3 shows the chronological distribution of compounds across centuries according to the OED earliest records attested:

	<i>cide</i>	<i>crat</i>	<i>ectomy</i>	<i>lith</i>	<i>logy</i>	<i>mania</i>	<i>morph</i>	<i>phile</i>	<i>phobia</i>	<i>scope</i>
14 <sup>th</sup>	1				4					
15 <sup>th</sup>	1				2					
16 <sup>th</sup>	4				13				1	
17 <sup>th</sup>	4				26			1		3
18 <sup>th</sup>		4			15	3			2	1
19 <sup>th</sup>	7	2	17	6	70	11	7	7	8	18
20 <sup>th</sup>	14	4	22	3	47	34	6	17	18	22
Listed	11	3	12	2	35	5	5	8	10	7
Not listed	3	1	10	1	12	29	1	9	8	15
	31	10	39	9	177	48	13	25	29	44

**Table 3: Chronological distribution of compounds across centuries by FCFs, according to the earliest record of each entry attested in the OED.**

#### 4 Combinatorial possibilities of final combining forms

The first aspect under examination is the status of the initial bases FCFs in the study sample combine with. The compounds in the sample were classified in terms of the morphological status of their initial bases as one of these three groups:

1. *Bound bases*. These are compounds where the first base cannot occur freely as an independent lexeme and does not have a free variant in English, e.g. *gerontocracy*, *lagomorph*, *stroboscope* or *xenolith*. This category also includes formations like *algicide*, where the initial base is a stem.<sup>9</sup> All bound bases are of classical origin, as shown in the examples. This means that, in this paper, *bound* correlates with the feature *classical origin*.
2. *Clipped bases*. These are compounds where the first base is formally bound in the compound but has a corresponding free morpheme in English, e.g. *Russophile* (*Russia[n]*), *colectomy* (*colon*), *Guggenmania* (*Guggenheim*) and *virology* (*virus*). Clipping occurs in various degrees in the study data, ranging from *virology* (*virus*) to *Eurocrat* (*Europe/European*).
3. *Free bases*. These are compounds where the first base stands as a free base in the compound, e.g. *kidneyectomy* (*kidney*), *oceanology* (*ocean*), *rodenticide* (*rodent*) or *colonoscope* (*colon*). The criterion for distinction between free and clipped bases is phonological, not orthographic. Therefore, formations like *virtuocracy* (*virtue*) are also in this group.

The definitions and the etymological information in the OED have been extensively used for the identification of the morphological status of the compounds' initial constituents. The analysis was synchronic and not etymological. Therefore, cases like *democracy* and *hydrophobia* are analysed as morphologically decomposable even if, according to the OED, they were borrowed into English as compound lexemes. Likewise, in cases where a clipped initial base has a free variant in English, the analysis does not take into consideration whether the free variant was in use or not in English at the time of the compound's formation. The general figures resulting from this classification are in Table 4:

<sup>9</sup> In this paper, *stem* is a bound lexical unit which takes a classical inflection.

	Bound		Clipped		Free		<i>p</i>
	Types	%	Types	%	Types	%	
<i>cide</i>	21	67.74	3	9.68	7	22.58	**
<i>crat</i>	7	70	1	10	2	20	n.s.
<i>ectomy</i>	13	33.33	9	23.08	17	43.59	n.s.
<i>lith</i>	9	100	0	0	0	0	**
<i>logy</i>	107	60.45	16	9.04	54	30.51	***
<i>mania</i>	14	29.17	5	10.42	29	60.42	**
<i>morph</i>	13	100	0	0	0	0	***
<i>phile</i>	14	56	5	20	6	24	n.s.
<i>phobia</i>	15	51.72	5	17.24	9	31.03	n.s.
<i>scope</i>	29	65.91	3	6.82	12	27.27	**
	242	56.94	47	11.06	136	32	***

**Table 4: Distribution of the compound types according to the morphological status of their initial base (bound, clipped or free).**

Results of goodness of fit and Fischer exact probability test are shown (*p* values, n.s. =  $p > .05$ ; \* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).<sup>10</sup>

As shown, over 50 % of the total compounds are formed on bound bases (56.94 %), and the distribution shows highly statistically significant differences. Table 4 also shows variation across the compounds formed on the various FCFs, even if the majority of the compound types also show a preference for bound bases. Indeed, most of them show percentages over 50 % in the category bound, while only two (*-ectomy* and *-mania*) show figures under 50 % in this category. This distribution is (highly) statistically significant in *-cide*, *-lith*, *-logy* *-morph* and *-scope* compounds for bound, and in *-mania* compounds, for free. No statistically significant differences are found in the distribution of *-crat*, *-ectomy*, *-phile* and *-phobia* compounds. A closer look at the compounds taking into consideration the OED earliest attestation dates brings to light relevant properties of the behaviour of these compounds. The types and percentages in the categories 'bound' and 'free' across the centuries are shown in Table 5:

	Bound		Free	
	Types	%	Types	%
14 <sup>th</sup>	5	100	0	0
15 <sup>th</sup>	2	66.66	1	33.33
16 <sup>th</sup>	15	83.33	3	16.66
17 <sup>th</sup>	28	82.35	6	17.64
18 <sup>th</sup>	20	80	3	12
19 <sup>th</sup>	113	74.34	26	17.10
20 <sup>th</sup>	59	31.21	98	51.85
Listed	44	89.79	43	43.87
Not listed	15	10.20	55	56.12

**Table 5: Diachronic distribution of the compounds according to the morphological status of their initial bases (free or bound).**

As observed, the percentages of compounds formed on bound initial bases decrease across the centuries, and they do steadily so from the 16<sup>th</sup> century onwards. This decrease correlates with

<sup>10</sup> Fischer exact probability test has been used in the study in cases where cells are <5. In Table 4 it has been used in the distributions of *-cide*, *-crat*, *-lith*, *-morph* and *-scope* compounds.

an increase in compounds formed on free initial bases.<sup>11</sup> In the 20<sup>th</sup> century, the tendencies become more marked and compounds formed on free initial bases account for over 50 % of the compounds, while compounds formed on bound initial bases cover only around one third of the compounds. In particular, as shown in Appendix I, bound bases in the 20<sup>th</sup> century account for 50 % of the formations in all compound types except for *-cide* (50 %) *-lith* (100 %) and *-morph* (100 %) compounds. Interestingly, also among 20<sup>th</sup> century compounds, those that are not listed in the OED tend to show the lowest percentages for boundness. This may indicate that non-formations, hapax legomena among them, tend to comprise compounds where the initial base is free. Exceptions to this are again *-lith* and *-morph*, and *-mania* and *-scope* compounds. However, this is probably because these compound types collect a low number of 20<sup>th</sup> century compounds (3 for *-lith*, 6 for *-morph*, 1 for *-mania* and 7 for *-scope* compounds). As a result, listed or non-listed 20<sup>th</sup> century compounds may appear as exceptionally high or low in percentages.

## 5 The linking element

For the analysis of the occurrence of a linking element, the formations were classified in two categories:

- 1 [-i/-o- + C], for compounds that contain the linking vowels *-i-* for the compounds ending in *-cide*, e.g. *insecticide*, and *-o-* for the compounds ending in any of the other FCFs in the study, e.g. *yankophile*.<sup>12</sup> The formations whose initial base ends in the expected linking vowel *-i-* or *-o-* (depending on the type of FCF, e.g. *suicide* and *mario-mania*, respectively) are also grouped here.
- 2 [C/V + C], for compounds that do not contain the linking vowels *-i-* or *-o-*, e.g. *cancerphobia*, *betjermania*. V stands for any vowel other than *-i-* in *-cide* formations, and *-o-* in the rest of the formations, e.g. *cinema-scope* and *olliermania*.

The compound type *-ectomy* has been excluded from the computations, because it is not subject to this feature. Phonological reasons may explain this, as it is the only FCF that starts in a vowel. The preceding segment is a consonant in the majority of cases, which is involved in *syllabification*. The two exceptions to these phonological combinatorial possibilities in the study sample are *kidneyectomy* and *myectomy* (Gr. *mys-*). Table 6 shows the general results.

---

<sup>11</sup> The small number of compounds formed in the 15<sup>th</sup> century (2 types), may have skewed a steady decrease from the 14<sup>th</sup> century. The same can be argued for the compounds formed on free initial bases in the 16<sup>th</sup> century (3 types) and in the 18<sup>th</sup> century (3 types), which could have also skewed the steady increase in the number of compounds formed on free initial bases.

<sup>12</sup> The compound *yank-o-phile*, illustrates the preference of FCFs for either *-o-* or *-i-* and, in particular, the preference of *-phile* compounds for *-o-* as a middle vowel: an *-i-* sound has been clipped from the original base (*Yan-kee*), which could have actually behaved as a middle vowel if *-i-* or *-o-* stood in free variation in the formation. Cf. however, *biocide*, *ethnocide* and *genocide* discussed below.

	<i>-i-, -o-</i> compounds	%	Total types	<i>p</i>
<i>cide</i>	29	93.55	31	***
<i>crat</i>	9	90	10	n.s.
<i>lith</i>	7	77.78	9	n.s.
<i>logy</i>	174	98.31	177	***
<i>mania</i>	23	47.92	48	n.s.
<i>morph</i>	11	84.62	13	n.s.
<i>phile</i>	24	96	25	***
<i>phobia</i>	21	72.41	29	n.s.
<i>scope</i>	31	70.45	44	n.s.
	329	85.23	386	***

**Table 6: Distribution of the compounds by occurrence of a linking vowel and by FCFs. Frequencies and percentages for each FCFs are shown.**

Results of goodness of fit and Fischer exact tests are shown (*p* values, n.s. = *p*>.05; \* = *p*<.05; \*\* = *p*<.01; \*\*\* = *p*<.001).

Overall, compounds containing a linking vowel amount to 85.23 % of the compounds. This distribution is highly statistically significant. By individual FCFs, most compounds also show very high percentages in the category *-i-, -o-* compounds. The differences in the distribution of the compounds are highly statistically significant in *-cide*, *-logy* and *-phile* compounds. No statistically significant differences have been found in the rest of compound types. Interestingly, the distribution of *-mania* compounds is not as uneven as in the previous cases: less than 50 % of *-mania* compounds (47.42 %) have a linking vowel. Table 7 shows the results according to the OED earliest attestation records:

	<i>cide</i>	<i>crat</i>	<i>lith</i>	<i>logy</i>	<i>mania</i>	<i>morph</i>	<i>phile</i>	<i>phobia</i>	<i>scope</i>	Types	%
14 <sup>th</sup>	100			75						4	80
15 <sup>th</sup>	100			100						3	100
16 <sup>th</sup>	100			100				100		18	100
17 <sup>th</sup>	100			92.3			100		66.66	31	91.18
18 <sup>th</sup>		100		100	100			100	100	25	100
19 <sup>th</sup>	100	100	75	100	90.9	85.71	100	87.5	94.44	130	95.59
20 <sup>th</sup>	71.42	75	100	100	29.41	100	87.5	61.11	59.09	118	71.52
Listed	70	100	100	74.46	20	83.33	61.11	81.81	30.76	75	87.21
Not listed	30	0	92.3	25.53	80	16.66	81.81	18.18	69.23	43	58.11

**Table 7: Diachronic distribution of the compound types containing a linking element as percentages. The sums of frequencies by centuries, and their respective percentages, are shown.**

As shown in the total results in the rightmost column of the table, the lowest percentage in the occurrence of a linking vowel occurs in the 20<sup>th</sup> century. This decrease is not as marked as the decrease in the occurrence of bound initial bases in the compounds described in the previous section. In particular, the only type of compounds showing figures under 50 % is *-mania* compounds (29.41 %).

The compounds lacking the expected linking vowel (*-i-* or *-o-* depending on the FCF) in *-cide*, *-crat*, *-lith*, *-logy*, *-morph* and *-phile* are: *biocide*, *ethnocide*, *genocide*, *virucide*, *Dixiecrat*,

*achnelith*, *megalith*, *genealogy*, *tetralogy*, *mineralogy*,<sup>13</sup> *polymorph*, *skeumorph* and *Detroit-phile*. Most of them are formed also on bound initial bases, which suggests that the lack of a linking element does not correlate with the occurrence of a free initial base. As suggested in the literature (cf. Bauer 1983: 214; Plag 2003: 158) and earlier in this paper, most compounds that do not incorporate the expected linking vowel contain initial bases ending in vowels. Relevant cases are *biocide*, *ethnocide*, *genocide* and *polymorph*, which do not show the expected linking vowel *-i-* for *-cide* compounds and *-i-* for *-morph* compounds. Arguably, in these cases the vowels they show are analysed as belonging to the left base. This can be seen in that they remain even if the final base starts with a vowel, as in *bioacoustics*, *ethnoarcheology* or *polyaxon*.<sup>14</sup> Finally, *virucide* and *skeumorph* also seem to preclude a linking vowel for the phonological reasons discussed in the previous cases. However, in this case we find that these two compounds are listed in the OED as *viricide* and *skeuomorph*, where the occurrence of (FCF-specific) linking vowels seems to override the phonological rule just discussed.

Some of the compounds formed on *-phobia* and *-scope* lacking a linking vowel also contain a bound initial base. They are *agoraphobia*, *acuphobia*, *hyperscope*, *stethoscope* and *telescope*. Again, the initial bases end in vowels in all cases (*hyperscope* only in non-rhotic accents). The rest of the compounds in this group lacking the default linking element are formed on free or clipped initial bases. Also, all *-mania* compounds lacking a linking element are formed on free initial bases. Table 8 shows the proportion of compounds formed on free and bound initial elements and lacking a linking element:

	Free initial base	%	Bound initial base	%
<i>mania</i>	22	88.00	0	0
<i>phobia</i>	5	62.5	2	25
<i>scope</i>	8	72.72	3	27.27

**Table 8: Distribution of *-mania*, *-phobia* and *-scope* compounds lacking the default linking elements into the categories free and bound initial bases.**

As shown, over 50 % of the compounds formed on free bases ending in *-mania*, *-phobia* and *-scope* also lack a linking vowel (compounds containing a bound initial base and lacking the linking element have been explained on phonological grounds in the previous paragraph). It seems significant that "free initial base" and "no linking vowel" should co-exist in *-mania*, *-phobia* and *-scope* compounds.

Finally, even if the majority of compounds without the default linking vowel combine with initial bases ending in vowels as we have seen in the previous paragraphs, there are also cases among compounds that lack linking vowels where the initial base ends in a consonant. This means that two consonants co-occur at the borderline. Interestingly, these cases are restricted to compounds formed on *-mania*, e.g. *wrestlemania*, *-phobia*, e.g. *child-phobia*, and *-scope*, e.g. *warp-scope*, where 56.52 % (*-mania*), 16.6 % (*-phobia*) and 28.57 % (*-scope*) of their compounds are formed on initial bases ending in consonants.<sup>15</sup>

<sup>13</sup> In *mineralogy*, there is reduction, or haplology, due to the similarity of the two syllables at the borderline of the two bases. Cf. however, *journalology*.

<sup>14</sup> No examples of *geno-* combining with a vowel initial base have been found.

<sup>15</sup> Compounds where the initial base ends in /r/ have been excluded from the computations to cater for rhotic pronunciations. This means that the percentages given would be higher if the computations catered for non-rhotic pronunciations.

## 6 Productivity

Two well-known complementary measures have been used to assess productivity: type frequency  $V$ , and productivity in the narrow sense (cf. Baayen/Lieber 1991). Type frequency refers to the number of units that follow a word-formation rule, and it is calculated by adding all the number of types containing the rule in question. The rationale of this measure is that the higher the  $V$  value is, the more productive the rule is considered to be. Computations based on types, although common in morphological studies, have been criticized for showing a plain and static picture of the activity of a morphological process and, in particular, for providing information only about past productivity (cf. Bauer 2001: 48–49; Plag 2005: 123–124). A more widely accepted computation of morphological productivity, and one that describes present productivity, is based on the *hapax legomena* in corpora, that is, corpus types with frequency 1. This measure assumes that the number of *hapax legomena* correlates with the number of neologisms and, therefore, that *hapaxes* are an indication of the extent to which a morphological rule produces new formations: the higher the number of *hapaxes* is, the higher the productivity of a morphological category is considered to be. Computations are according to the formula below, where  $n_1$  is the number of *hapaxes* containing a word-formation rule, and  $N$  is the total number of tokens with that rule. The results are between 0 and 1, where 1 signals the most productive rule.

$$P = \frac{n_1}{N}$$

This paper aims to explore changes in the productivity of FCFs, so the choice of these two measures seems particularly suitable to show patterns of past productivity (type frequency  $V$ ), and of current productivity ( $P$  productivity). Table 9 shows the results for  $V$  and  $P$  productivity.

	$V$	$P$
<i>logy</i>	177	0.0010
<i>mania</i>	48	0.2459
<i>scope</i>	44	0.0111
<i>ectomy</i>	39	0.0172
<i>cide</i>	31	0.0045
<i>phobia</i>	29	0.0389
<i>phile</i>	25	0.1417
<i>morph</i>	13	0.0833
<i>crat</i>	10	0.0030
<i>lith</i>	9	0.0217

**Table 9: Results according to  $V$  and  $P$  productivity by FCFs, ranked by highest  $V$  values.**

Table 9 shows that there are major productivity differences from one FCF to another, regardless of the productivity measure. Interestingly, both measures place one FCF as the most productive FCF in the study sample by large. For  $V$ , it is *-logy* which is over three times more productive than the second and third most productive FCFs (*-mania* and *-scope*), and about twenty times more productive than the lowest two FCFs (*-crat* and *-lith*). By contrast,  $P$  ranks *-mania* as the most productive FCF, again about twice more productive than the second highest FCF in the rank (*-phile*), and at a considerable distance from the lowest two FCFs in the ranking of  $P$  productivity (*-crat* and *-logy*), which are over twenty times less productive.

Each of the measures also gives different productivity values for one and the same FCF. Figure 1 displays the productivity ranking for each FCF obtained from each computation. Kendall's tau  $\tau$  test, which measures the similarity between two ranked sets of quantified items, confirms the divergence of results from each measure ( $p=0.7205$ ). A case in point is *-logy*, which ranks highest according to type frequency but lowest according to  $P$ . Other major ranking differences

are in *-lith*, *-morph* and *-phile*. Small differences can still be found, marked in grey in Figure 1: *-mania* is the most productive FCF according to *P*, and the second most so according to type frequency; in *-crat* both measures converge and it therefore ranks ninth highest, marked in black.

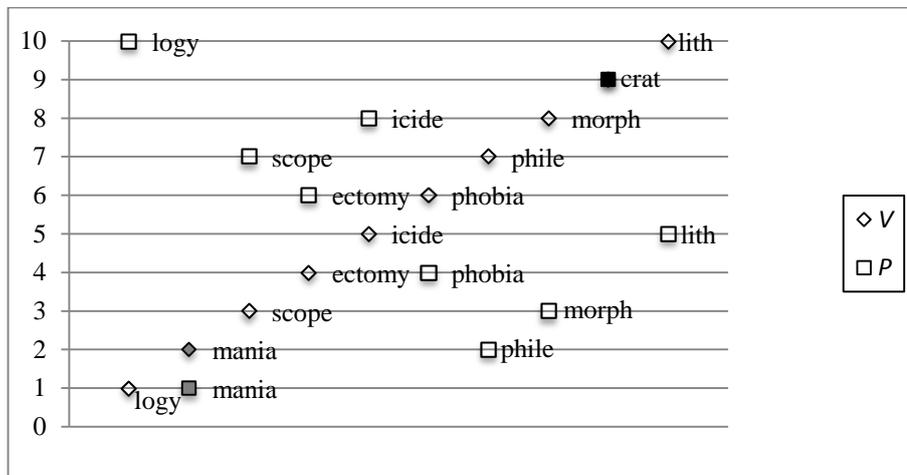


Figure 1: Ranking of the FCFs according to *V* and *P* productivity.

Discrepancies in the results given by the two measures have been anticipated in the first part of this section (cf. also Bauer 2001: 48–49; Plag 2005: 123–124). They are explained in that each measure captures different aspects of productivity: *V* represents past productivity and *P* represents the potential of a rule to produce new coinages. This discrepancy also suggests that productivity of units changes over time and, for example, that, while *-logy* has produced by far the largest number of compounds in the study, at the point of time represented by the corpus (1980s–1993), it does not produce as many new constructions as *-mania*.

This hypothesis has been further investigated by examining the proportion of formations first recorded in the 20<sup>th</sup> century for each FCF, which presumably will give indications of each FCF's productivity in most recent times with respect to previous chronological stages. Formations whose listing date could not be attested in the OED have been this time grouped with listed 20<sup>th</sup> century units. The results are in Figure 2:

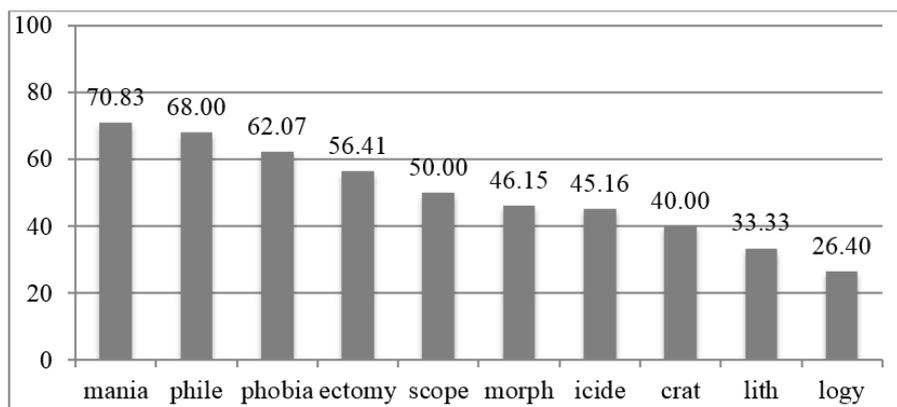
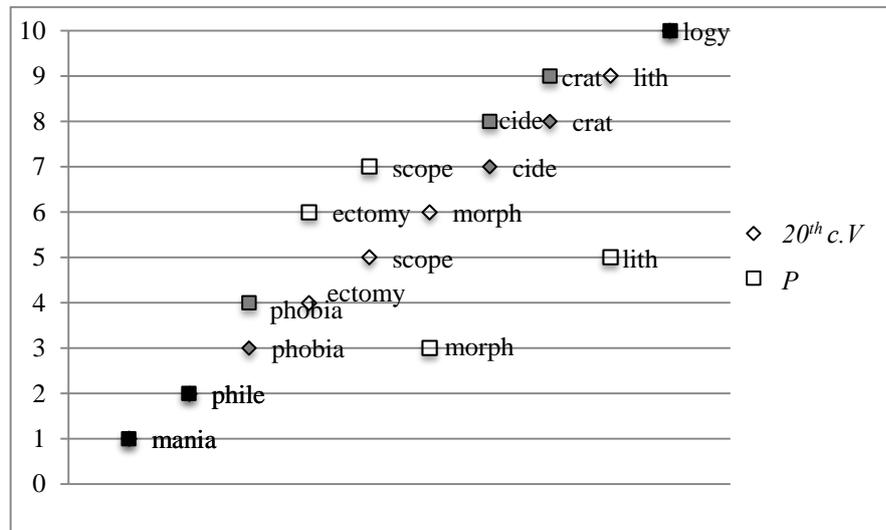


Figure 2: Percentages of types by FCFs dating from the 20<sup>th</sup> century, arranged from most to least productive FCFs.

Figure 2 shows, first, that the ranking of compounds is now similar to that obtained from *P* above (Kendall's tau  $\tau$  test,  $p=.007$ ). This is shown more clearly in Figure 3 below, which displays the ranking of the type frequency proportions and *P*. This time *-mania* and *-phile* rank as the two most productive FCFs and *-logy* as the least productive one. Also notably, the rest of

the FCFs for both measures seem to stand very close to their counterparts, with the exception of *-morph* and *-lith*:



**Figure 3: Ranking of compounds by FCFs according to V for 20<sup>th</sup> century compounds and P. Darker colour means greater correlation of the rankings.**

In addition to the similarity of rankings, a second interesting aspect of Figure 2 is that at least 50 % of the compounds ending in *-mania*, *-phile*, *-phobia*, *-ectomy* and *-scope* date from the 20<sup>th</sup> century. A third point about Figure 2 is that *-logy*, which ranked highest in the formation of compounds in overall type frequency computations, now ranks lowest (only 26.40 % of its compounds are formed during the 20<sup>th</sup> century). These two latter observations may be explained in that most of the compounds in our study formed on *-mania*, *-phile*, *-phobia*, *-ectomy* and *-scope* date from the 19<sup>th</sup> century onwards, while *-logy* compounds in our study date from as early as the 14<sup>th</sup> century. This means that *-logy* has been active for a larger period of time and has produced many more forms than the other compounds, but apparently it is not as productive as other FCFs today. Table 10 shows the distribution of compounds across the centuries they date from:

	<i>cide</i>	<i>crat</i>	<i>ectomy</i>	<i>lith</i>	<i>logy</i>	<i>mania</i>	<i>morph</i>	<i>phile</i>	<i>phobia</i>	<i>scope</i>
14 <sup>th</sup>					2.26					
15 <sup>th</sup>					1.13					
16 <sup>th</sup>					7.34				3.45	
17 <sup>th</sup>					14.69			4		6.82
18 <sup>th</sup>		40			8.47	6.25			6.90	2.27
19 <sup>th</sup>	22.58	20	43.59	66.67	39.55	22.92	53.85	28	27.59	40.91
20 <sup>th</sup>	45.16	40	56.41	33.33	26.55	70.83	46.15	68	62.07	50
Listed	78.57	75	54.55	66.67	74.47	14.71	83.33	47.06	55.56	31.82
Not listed	21.43	25	45.45	33.33	25.53	85.29	16.67	52.94	44.44	68.18

**Table 10: Diachronic distribution of compounds by FCFs as percentages. Cf. Table 3 for frequencies.**

As can be seen, most compounds in the study sample date from the 19<sup>th</sup> and 20<sup>th</sup> centuries. In fact, the OED lists most compounds among earlier centuries as compounds borrowed directly from classical Greek or Latin, e.g. *philology* (14<sup>th</sup> century), *astrology* (15<sup>th</sup> century), *hydrophobia* (16<sup>th</sup> century), *infanticide* (17<sup>th</sup> century) or *democrat* (18<sup>th</sup> century), which means that they are not English formations.

Also, according to Table 10, most FCFs show increasing figures from the 19<sup>th</sup> century onwards, notably *-cide*, *-mania*, *-phile*, *-phobia* and *-scope*. Others, e.g. *-lith* and *-morph* and, most remarkably, *-logy*, show the opposite tendency. Finally, over 50 % of the formations in the 20<sup>th</sup> century *-mania*, *-phile* and *-scope* are contributed by the compounds in the 'Not listed' row, i.e. by compounds in the corpus that are not listed in the OED. These are often hapaxes, which may partly explain why *-mania* and *-phile* rank as the most productive FCFs for *P*.

## 7 Discussion

### 7.1 General appraisal of the findings

From the results described above, the following conclusions can be drawn. Regarding the morphological status of the initial element, and even if a large number of compounds are formed on bound initial bases, current tendencies show that selection of bound bases is no longer the default choice. The results show that most compounds on bound initial bases were formed in earlier dates, while compounds on free bases date mainly from the 19<sup>th</sup> and 20<sup>th</sup> centuries. To give an example, at least 50 % of the 20<sup>th</sup> century compounds ending in *-crat*, *-ectomy*, *-logy*, *-mania* and *-scope* are formed on free initial bases. Exceptions are *-lith* and *-morph* compounds, where the feature 'bound' remains constant across all their compounds.

By contrast, an exploration of the occurrence of the linking element reveals a marked occurrence of the linking vowel for all the compound types in overall computations and also diachronically, except for *-mania*.<sup>16</sup> Most of the compounds without a linking element can be explained in that they are formed on initial bases ending in vowels. This finding further supports the claims that the occurrence of linking vowels has the phonological function of preventing two consonants from co-occurring at the borderline (cf. Bauer/Lieber/Plag 2013: 456). The FCF *-mania* is the only FCF where less than 50 % of its compounds in the 20<sup>th</sup> century contain a linking element (29.41 %). The next two compound types in terms of frequency, even though with frequencies of over 50 %, are *-scope* (59.09 %) and *-phobia* (61.11 %) compounds. Interestingly, *-mania*, *-scope* and *-phobia* are the exceptions in the phonological constraint just mentioned for the occurrence of a linking vowel, that is, they gather all the cases where two consonants co-occur at the borderline.

Regarding productivity, the findings show that not all the FCFs in the study are equally productive, and also that the productivity of the FCFs also varies over time. According to *P* productivity, *-mania*, *-phile* and *-phobia* rank as the most productive FCFs nowadays, while *-logy*, *-crat* and *-cide* rank as the least productive. Over time, all FCFs, except *-lith*, *-logy* and *-morph*, show increasing figures towards the 20<sup>th</sup> century. This increase is particularly marked in *-mania*, *-phile*, *-phobia* and *-ectomy* compound types, which form at least 50 % of the compounds in the 20<sup>th</sup> century. *P* ranks *-logy* as the least productive FCF in the group, which is consistent with the decreasing number of types registered towards the 20<sup>th</sup> century. However, *P* also ranks *-lith* and *-morph* as the fifth and third most productive units respectively, which is not consistent with the decreasing tendency that these compound types show towards the 20<sup>th</sup> century according to *P*. This inconsistency may be due to the low number of *-lith* and *-morph* compounds in the study.

The findings confirm that FCFs are far from a homogeneous class. Still, they also show that some of them are more similar than others. The FCFs *-crat*, *-lith* and *-morph* stand among the most prototypical FCFs in their morphological behaviour, as most of their compounds are formed on bound bases and show a linking element. Notably, they are also among the least productive CFs in the study sample. One third of the compounds formed on these FCFs are also borrowed compounds. By contrast, *-mania*, *-phile*, *-phobia* and *-scope* stand among the least

<sup>16</sup> *-ectomy* was not submitted to analysis for the phonological reasons explained above.

prototypical FCFs. They do not show a clear tendency towards bound initial bases, and *-mania*, *-phobia* and *-scope* gather compounds that show a distinct behaviour regarding the occurrence of the linking element. The FCFs *-mania*, *-phile* and *-phobia* also stand among the most productive FCFs nowadays. The remaining FCFs (*-logy*, *-ectomy* and *-cide*) behave dissimilarly depending on the aspect under examination.

## 7.2 Morphological implications of the findings

The data presented in the paper seems to show that, as neoclassical combining elements increase in productivity, they become more absorbed into native patterns of word-formation: an increase in the number of formations towards the 20<sup>th</sup> century, occurs in combination with an increase in the co-occurrence of the FCFs with free initial bases. The co-existence of these two tendencies may be taken as an indication of a suffix-like status for terminal units like *-logy*, *-phile* or *-ectomy*. However, this statement becomes controversial in a synchronic analysis of bound terminal units in neoclassical compounds where, along examples like *Egyptology*, there are formations like *neology* that would preclude an analysis of *-logy* as a suffix (p.c. with Laurie Bauer 2013). Therefore, the statement that some FCFs in the study are better analysed as suffixes will hold if we accept that their status is evolving.

Compounds in *-mania* also show high present and past productivity and a marked preference for free initial bases or no particular preference for any of the types of initial bases considered in the study. Other similar cases may be *-phobia*, and probably also *-scope*. However, the morphological implications here may be different given that they also show a number of distinct features. All three can stand as free morphemes, *scope* possibly as a shortened form of *microscope* or *telescope*. In addition, they do not always show a (linking) vowel between the two morphemes, allowing the occurrence of two consecutive consonants, e.g. *Beatlemania*, *stormscope*. The latter examples co-exist with others like *kleptomania*, *ailurophobia* and *endoscope*, where the FCFs combine with bound bases and a linking element surfaces. In this particular situation, however, it seems that two types of compounds on *-mania*, *-phobia* and *-scope* exist, one of them neoclassical (*kleptomania*, *ailurophobia*) and the other native (*Beatlemania*, *stormscope*). As pointed out also for *-itis* (cf. section 4), use/meaning extension may be one crucial aspect in this distinction, as *-mania* and *-phobia* cover in the study sample both medical (*pyromania*, *photophobia*) and also non-medical uses (*Beatlemania*, *cancerphobia*). To what extent this distinction can be associated with the diverging morphological behaviours is to be explored.

Finally, as shown by the findings, and except for the cases discussed above, it appears that one of the most characteristic properties of neoclassical compounds is the occurrence of a linking vowel. This remains a marked feature in the 20<sup>th</sup> century. In cases where this feature co-occurs with FCFs that combine mainly with free initial bases, which is the case of *-logy* and *-phile*, a listing of these FCFs as *-ology*, and *-ophile* is to be supported. Incidentally, the resulting phonological configuration of these bound terminal units –vowel initial– is another aspect which they happen to share with neoclassical suffixes.

## 8 Conclusion

Synchronically, the class of neoclassical compounds proves to be as heterogeneous as predicted in the literature. Diachronically, earlier compounds are more prototypical in their behaviour, while more recent formations tend to be less so. Recent formations seem to show features of native patterns of word-formation, namely the co-occurrence of bound terminal units with free initial bases. In turn, this places terminal units halfway between bases and suffixes. The occurrence of the linking element is in most cases a constant feature across neoclassical compounds and it is so over time too. The constant occurrence of a linking vowel before terminal units which attach to free bases makes endings like *-ology* and *-ophile* appear to be even more suffix-like.

Among these generalisations, the case of *-mania* and *-phobia* stand out as the least prototypical type of compounds in the study in various respects. Whether, at least in some particular uses, they should stand as elements involved in native compounding or not, needs further consideration.

## References

- Adams, Valerie (2001): *Complex Words in English*. Harlow: Pearson.
- Allen, Margaret (1978): *Morphological Investigations*. Ann Arbor/MI: University of Connecticut. Dissertation.
- Amiot, Dany/Dal, Georgette (2007): "Integrating Neoclassical Combining Forms into a Lexeme-based Morphology". *Online Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*: 322–336. <http://w3.erss.univ-tlse2.fr:8080/index.jsp?perso=lasserre&subURL=Actes.pdf>, accessed October 14, 2014.
- Aronoff, Mark/Fuhrhop, Nanna (2002): "Restricting Suffix Combinations in German and English. Closing Suffixes and the Monosuffix constraint". *Natural Language and Linguistic Theory* 20/3: 451–490.
- Baayen, Harald/Lieber, Rochelle (1991): "Productivity and English Derivation. A Corpus-based Study". *Linguistics* 29: 801–844.
- Baeskow, Heike (2004): *Lexical Properties of Selected Non-native Morphemes of English*. Tübingen: Narr.
- Bauer, Laurie (1983): *English Word-formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie (1998): "Is there a Class of Neoclassical Compounds, and if so is it Productive?" *Linguistics* 36: 403–422.
- Bauer, Laurie (2001): *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie/Huddleston, Rodney (2002): "Lexical Word-formation". In: Huddleston, Rodney/Pullum, Geoffrey K. (eds.): *The Cambridge Grammar of the English Language*. Cambridge, Cambridge University Press: 1621–1721.
- Bauer, Laurie/Lieber, Rochelle/Plag, Ingo (2013): *Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Booij, Geert (1992): "Compounding in Dutch". *Rivista di Linguistica* 4: 37–59.
- Brigham Young University (ed.): *British National Corpus*. <http://corpus.byu.edu/bnc/>, accessed October 14, 2014.
- Donalies, Elke (2009): "Stiefliches Geofaszintainment. Über Konfixtheorien". In: Müller, Peter O. (ed.): *Studien zur Fremdwortbildung*. Hildesheim, Olms: 41–64. (= *Germanistische Linguistik* 197/198).
- Giegerich, Heinz J. (1999): *Lexical Strata in English. Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.
- Iacobini, Claudio (2000): "Base and Direction of Derivation". In: Booij, Geert/Lehmann, Christian/Mugdan, Joachim (eds.): *Morphologie/Morphology. Ein internationales Handbuch zur Flexion und Wortbildung/An International Handbook on Inflection and Word-Formation*. 1. Halbband/Volume I. Berlin/New York, de Gruyter: 865–876.
- Iacobini, Claudio (2004): "Composizione con Elementi Neoclassici". In: Grossmann, Maria/Rainer, Franz (eds.): *La Formazione delle Parole in Italiano*. Tübingen, Max Niemeyer: 69–95.
- Kastovsky, Dieter (2009): "Astronaut, Astrology, Astrophysics. About Combining Forms, Classical Compounds and Affixoids". In: McConchie, Roderick/Honkapohja, Alpo/Tyrkkö, Jukka (eds.): *Selected Proceedings of the 2008 Symposium on New Approaches in English Historical Lexis (HEL-LEX 2)*. Somerville/MA, Cascadilla Proceedings Project: 1–13.
- Kilgarriff, Adam (1996): *Unlemmatised Full BNC List*. <http://www.kilgarriff.co.uk/BNCLists/all.al.o5>, accessed April 07, 2013.

- Kiparsky, Paul (1982): "Lexical Morphology and Phonology". In: Yang, In-Seok (ed.): *Linguistics in the Morning Calm*. Seoul, Hanshin: 3–91.
- Kirkness, Alan (1995): "Eurolatin – the Greek and Latin Patrimony in the European Languages". In: Institut für Deutsche Sprache (ed.): *Lexicographica 11*. Tübingen, Niemeyer: 262–265.
- Lüdeling, Anke/ Evert, Stefan (2005): "The Emergence of Productive Non-medical *-itis*. Corpus Evidence and Qualitative Analysis". In: Kepsner, Stephan/Reis, Marga (eds.): *Linguistic Evidence. Empirical, Theoretical, and Computational Perspectives*. Berlin, Mouton de Gruyter: 351–370.
- Lüdeling, Anke/Schmid, Tanja/Kiokpasoglou, Sawwas (2002): "Neoclassical Word Formation in German". In: Booij, Geert/Marle, Jaap van (eds.): *Yearbook of Morphology 2001*. Berlin, Springer Netherland: 253–283.
- Marle, Jaap van (1985): *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.
- Oxford English Dictionary Online* (2014). Oxford: Oxford University Press.
- Plag, Ingo (2003): *Word-formation in English*. Cambridge: Cambridge University Press.
- Plag, Ingo (2005): "Productivity". In: Brown, Keith (ed.): *Encyclopedia of Language and Linguistics*. 2<sup>nd</sup> ed., vol. 10. Oxford, Elsevier: 121–128.
- Prčić, Tvrtko (2008): "Suffixes vs. Final Combining Forms in English. A Lexicographic Perspective". *International Journal of Lexicography* 21/1: 1–22.
- Siegel, Dorothy (1974): *Topics in English Morphology*. Cambridge/MA: MIT. Dissertation.
- Stanforth, Anthony W. (2005): "Effects of Language Contact on the Vocabulary. An Overview". In: Cruse, D. Alan et al. (eds.): *Lexikologie/Lexicology*. Berlin/New York, de Gruyter: 805–813.
- Warren, Beatrice (1990): "The Importance of Combining Forms". In: Dressler, Wolfgang U. et al. (eds.): *Contemporary Morphology*. Berlin, Mouton de Gruyter: 111–132.

**Appendix I. Diachronic distribution of the compound types in the study sample according to the morphological status of their initial bases (bound or free). Percentages are shown.**

	<i>cide</i>		<i>crat</i>		<i>ectomy</i>		<i>lith</i>		<i>logy</i>		<i>mania</i>		<i>morph</i>		<i>phile</i>		<i>phobia</i>		<i>scope</i>	
	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free	Bound	Free
14 <sup>th</sup>	100	0				100	0													
15 <sup>th</sup>	100	0				50	50													
16 <sup>th</sup>	100	0				76.92	23.08								100	0				
17 <sup>th</sup>	75	25				80.77	19.23						100	0						
18 <sup>th</sup>			100	0		66.66	20	100	0						100	0				
19 <sup>th</sup>	71.43	28.57	100	0	52.94	35.29	100	0	65.21	21.73	90.91	9.09	100	0	71.43	14.28	75	12.5	100	0
20 <sup>th</sup>	50	28.57	25	50	18.18	50	100	0	33.33	56.25	2.94	82.35	100	0	47.05	29.41	33.33	44.44	31.81	54.54
Listed	85.71	100	100	50	75	72.72	66.66	0	93.75	62.96	0	17.85	83.33	0	62.50	20	83.33	25	28.57	38.46
Not listed	14.28	0	0	50	25	27.27	33.33	0	6.25	37.03	100	82.14	16.66	0	37.50	80	16.66	75	71.42	61.53