

Article

A Construction Morphology Approach to Neoclassical Compounds and the Function of the Linking Vowel

Hiromi Hayashi

Department of Pharmacy and Health Sciences, Meiji Pharmaceutical University, Tokyo 204-8588, Japan; hhayashi@my-pharm.ac.jp

Abstract: The morphological status of combining-forms (CF) (*bio-*, *-logy*, etc.) used in neoclassical compounds (*biology*, *bioscience*, etc.) is a matter of debate in morphology. Some see them as having the same status as ordinary words, while others see them as having a special status and forming an independent category. This paper focuses on neoclassical compounds having basic and extended patterns. The former, *biology* type, is composed solely of CFs, while the latter contains ordinary words, i.e., free morphemes (FM), and is further classified into two categories: *bioscience* type and *hamburgerology* type. This paper aims to explain the relationship among those subtypes of neoclassical compounds while capturing their extension processes within the framework of Construction Morphology (CxM). In particular, we outline a novel analysis of those subtypes of neoclassical compounds using “schemas” and “constructional idioms” following the literature regarding CxM.

Keywords: neoclassical compound; combining form; linking vowel; construction; schema; constructional idiom



Citation: Hayashi, Hiromi. 2024. A Construction Morphology Approach to Neoclassical Compounds and the Function of the Linking Vowel. *Languages* 9: 129. <https://doi.org/10.3390/languages9040129>

Academic Editors: Akiko Nagano and Ryohei Naya

Received: 14 December 2023

Revised: 8 March 2024

Accepted: 10 March 2024

Published: 1 April 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the features of neoclassical compounds (*biology*, *bioscience*, etc.) is that they contain combining forms (CFs) (*bio-*, *-logy*, etc.), which are bound morphemes that typically originate from classical languages such as Greek and Latin. Researchers have different views on what a combining form is. First, there is the position that combining forms are lexemes, as with typical words like *book* and *science*. In the words of Nagano and Shimada (2014, p. 330), “lexemes are abstract units with a rich internal structure that binds form, grammar, and meaning together”. A lexeme may be realized as a free morpheme or a bound form, and the constituents of the neoclassical compound (i.e., *bio-*, *-logy*, etc.) are considered bound forms in this sense. For example, *biology* and *bioscience* should be formed by combining *bio-*, *-logy*, and *science*, all of which are “lexemes”. Then, neoclassical compounds are formed by combining lexemes in the same way as ordinary compounds.

Another position is that, morphologically, there is a unique category covering CFs. By a “unique category”, we mean that CFs are not lexemes; if a compound is formed by combining lexemes, then a neoclassical compound consisting of CFs, which are not lexemes, is not a compound.

Therefore, from the viewpoint of morphology, it makes a big difference whether a CF is a lexeme or not. In the former case, a neoclassical compound is a “compound” formed by compounding; in the latter case, a neoclassical compound is a “non-compound” word.

In this paper, we pay great attention to the fact that some neoclassical compounds, such as *bioscience* and *hamburgerology*, contain free morphemes (FMs), in addition to CFs. Then, there are three subtypes of neoclassical compounds:

- (1) a. CF-CF
- b. CF-FM
- c. FM-CF

In (1), a linking vowel (LV), which is usually realized as *-o-*, occurs between CFs, or at the linkage between CFs and FMs, as in *bi-o-logy*, *bi-o-science*, and *hamburger-o-logy*. It is explicitly shown in (2) as follows:

- (2) a. CF-*o*-CF
 b. CF-*o*-FM
 c. FM-*o*-CF

If we discard the distinction between CF and FM in (2), we can say that a neoclassical compound has the following form (3):

- (3) X-*o*-Y

Thus, neoclassical compounds exhibit a major formal feature that CF presents in both or one of the positions X and Y, and LV appears in the linking part.

In this paper, we argue that it is necessary to treat neoclassical compounds differently from ordinary compounds since neoclassical compounds have some other characteristics that are different from ordinary compounds in English, such as the presence of LV, and are often used in academic fields. We treat them as expressions with the “form”, or “schema”, represented in (3) above. Construction Morphology (CxM) (e.g., [Booij 2010](#); [Booij and Hüning 2014](#); [Arcodia and Basciano 2018](#); [Booij and Audring 2018](#)) is a useful framework that allows for such an analysis. Therefore, in this paper, we discuss neoclassical compounds in the framework of CxM and explain their specific properties.

In addition, we propose that the hierarchical relationships among schemas, constructional idioms, and individual words, as well as their interactions, explain the diversity and productivity of the neoclassical compounds. Thus, we claim that the CF does not have independent status in morphology, and it is merely a name of convenience. Our assumption is that what has been referred to as the CF is the element used in the schema ‘X-*o*-Y’, which is a generalized schema for the neoclassical compounds with variables to be filled with CFs or FMs and a constant *-o-*.

This paper is organized as follows: Section 2 discusses the morphological status of CF, Section 3 focuses on the status of LV, and Section 4 analyzes the neoclassical compounds in CxM, with particular attention to constructional idioms, one of the tools used in CxM. Section 5 presents a summary of the findings and future perspectives.

2. The Status of Combining Forms

Section 2.1 discusses previous studies that analyze CF as a lexeme, while Section 2.2 discusses those that analyze CF as a special word-formation element other than a lexeme.

2.1. CF-as-Lexeme Analysis

[Kastovsky \(2009a, 2009b\)](#) argues that CF is a type of lexeme and that neoclassical compounds are, therefore, compounds. According to this view, neoclassical compounds are not a special form; for example, words shown in (4) are considered stem compounds formed from stems, which are the realization of lexemes:

- (4) astr-o-naut, patr-i-cide, galvan-o-scope ([Kastovsky 2009b](#), pp. 1, 10)

A stem is a bound form that is also used in derivation, like *scient-* and *dramat-*, as shown in (5), and does not appear only in neoclassical compounds:

- (5) scient_N-(-ist) vs. science, dram_{tN}-(-ic) vs. drama ([Kastovsky 2009b](#), p. 9)

In other words, there is no need for a special concept of CF; studies that share [Kastovsky’s](#) view include [Adams \(1973, 2001\)](#) and [Yoshioka \(2011\)](#).

In [Shimada and Nagano \(2011a, 2011b\)](#) and [Nagano and Shimada \(2014\)](#), neoclassical compounds are thought to be formed from lexemes in the same way as ordinary com-

pounds. For example, both *psycho-*, which is CF, and *mind*, which is FM, are morphological realizations of the same lexeme MIND. Similarly, *-logy* and *study* are morphological realizations of the same lexeme STUDY. Then, *psychology* and *mind study* are both created by combining the constituents in the same way. (cf. Nagano and Shimada (2014, p. 357), Shimada and Nagano (2011a)). The difference between ordinary compounds such as *mind study* and neoclassical compounds such as *psychology* is whether the realization of lexemes consists solely of FMs or contains bound forms.

However, if we assume that neoclassical compounds are formed from lexemes, then we must consider a special mechanism to deal with LVs that occur in the linking parts of neoclassical compounds, such as *-o-* in *astronaut* and *-i-* in *patricide*. LV occurrence is a major difference between neoclassical and ordinary compounds, and their occurrence must be explained in some way.

One more analysis is that of Baeskow (2004), who, while identifying CFs as lexemes, establishes a mechanism specific to neoclassical compounds: a typical structure of neoclassical compounds proposed in Baeskow (2004) is shown in (6).

- (6) neoclassical configurations
- | | | |
|-----------------|---------------------------------|-----------------------|
| nc-structure 1: | [[<u> </u> / <u>V</u>]/][X]] | |
| nc-structure 2: | [[X]/ <u>V</u>]/[<u> </u>]] | |
| X = | [(bound) root, -Germanic]] | |
| /V/ = | /O/ | (Baeskow 2004, p. 99) |

In this analysis, the constituents in the neoclassical compound are bound roots, which occur in the blank slots of (6). For example, *phon/phone*, as shown in (7), occurs in the blank slot of nc-structure 1 when it is used as the initial part, as in *phonology*, and in the blank slot of nc-structure 2 when it is used as the final part, as in *telephone*. (*nc* is an abbreviation of ‘neoclassical compound’ (Baeskow 2004, p. 98 fn 48).)

- (7) Lexical entry for the bound root *phon*
 orthographic representations: phon, phone
 phonological representation: /fOn/

formal features:

| |
|--|
| bound root +common +count +concrete -human <R> -Germanic |
|--|

configuration frames: \diamond nc-structures 1 and 2
 [[] X]; X=-al, -ate, -ic



ϵ -feature: [Greek]
 orthographic representation: $\phi\omega\nu\acute{\eta}$
 semantic representation: [VOICE]

(Baeskow 2004, pp. 100–1)

(7) also specifies that *phon/phone* occurs with the suffixes *-al*, *-ate*, and *-ic*, making it clear that the constituents of the neoclassical compounds are not exclusively used in

this construction. (“Etymological Component” is reserved for those with etymological knowledge.)

Baeskow’s (2004) analysis, while assuming that CF is a bound root, i.e., a lexeme, proposes an operation other than compounding for neoclassical compounds formed from lexemes. Furthermore, in distinguishing between the levels of compound formation because the study is within the framework of the Minimalist Program, Baeskow (2004, p. 85) proposes that the preceding CFs such as *micro-* and *bio-* are reinterpreted as class II prefixes only when they select free morphemes at level 2, forming neoclassical compounds such as *microwave* and *biofeedback*. These special operations that cannot be used for anything other than neoclassical compounding are undesirable from a theoretical point of view.

2.2. CF-with-Specific-Status Analysis

Let us look at Bauer (1983), a pioneering study of neoclassical compounds in English, which takes the position that CF is neither a lexeme nor an affix but a specific word-formation element.

Looking at the CFs used in the neoclassical compounds *biology* and *electrophile*, we see that some, such as *bio-* and *electro-*, occur at the beginning of the word, and some, such as *-logy* and *-phile*, occur at the end of the word. Bauer distinguishes between those types, calling the former Initial Combining Form (ICF) and the latter Final Combining Form (FCF). Importantly, lexemes, which are the building blocks of ordinary compounds, do not have positional restrictions such as “occurring at the beginning of the word” or “occurring at the end of the word”. For example, lexemes like *life* can occur at both the beginning and the end of words: *life science*, *lifestyle*; *research life*, *school life*. In contrast, combining forms such as *bio-* only occur at the beginning of the word, as in *biology* and *bioscience*. Thus, under this analysis, CFs are different from the constituents of ordinary compounds. Furthermore, Bauer argues that ICF and FCF are also different from affixes. The reason for this is that if *electro-* and *-phile* were affixes, then *electrophile* would be a word composed entirely of affixes, which would violate the condition that affixes usually occur with word bases (including word roots) (cf. Bauer 1983, pp. 213–16). In conclusion, in Bauer’s view, CF is neither a lexeme nor an affix.

Other analyses of CF as a special word-formation element include Warren (1990) and Prčić (2005, 2008). They follow Bauer (1983) in dividing CFs into ICFs and FCFs according to their position, and they also focus on their origin.

In Prčić’s study (2005, 2008), CFs are classified into the four types shown in (8) according to the positions and the origins of the CFs. In (8), the italicized parts are classified as ICFs and FCFs.

- (8) a. classical ICF: *xylo-*(-phone), *calli-*(-graphy), *alti-*(-meter)
 b. modern ICF: *jazzo-*(-phile), *speedo-*(-meter), *heli-*(-port) (a-b, Prčić 2005, p. 317)
- c. classical FCF: (morpho-)-*logy*, (demo-)-*cracy*, (biblio-)-*phile*
 d. modern FCF: (coffee-)-*holic*, (sugar-)-*holic*; (six-)-*footer*, (fifth-)-*grader*,
 (blue-ey)-*ed*; (scandal-)-*monger*, (gab-)-*fest*, (horror-)-*meister*, (dulls-)-*ville*
 (c-d, Prčić 2008, pp. 5–7)

One of the remarkable assumptions in Prčić’s (2005) system is the explicitly generalized form of (9), i.e., “Modern English word + linking vowel”, which creates modern ICFs in (8b).

- (9) FM + -o-

By using (9), a modern ICF is created, including the form of “proper noun + -o-”, such as *Egypto-* in *Egyptology*. Then, “modern ICFs need not be entered in dictionaries because their meaning, related to the base word, is self-explanatory, and their -o- ending is largely predictable, too” (Prčić 2005, p. 317). It is noteworthy that the introduction of a mechanism to create a CF by attaching -o- to FM explains the specific nature of CF and its high productivity while regarding CF as a special word formation element.

Among the previous studies that regard CFs as a unique category, [Amiot and Dugas \(2020\)](#) is also important; [Amiot and Dugas \(2020, p. 859\)](#) argues that “CFs (i) are bound forms, (ii) semantically correspond to lexemes..., (iii) do not serve to form large series”. While (i) and (ii) are similar claims to the previous studies we have seen, (iii) is not found in any other previous studies. By “do not serve to form large series”, Amiot and Dugas mean that CF should not be involved in the formation of many words. Thus, the restriction (iii) excludes *macro-*, *pseudo-*, *-logist*, *-cide*, *-vorous*, etc., which have been conventionally treated as CFs. In their argument, they are not CFs because they are involved in the formation of many words but rather affixes or affixoids, which are content words that act like affixes (cf. [Ralli 2013](#), pp. 228–29, among others).

It should be noted that a common problem with arguments that treat CF as its own category is that the distinction between CFs and affixes is not clear, or they are treated differently by different researchers. For example, [Warren \(1990, p. 123\)](#) notes that *neo-* is considered a prefix in [Marchand \(1969, pp. 178–79\)](#), while [Quirk et al. \(1985, pp. 1545–46\)](#) treat it as a CF. Conversely, in the case of *pseudo-*, [Marchand \(1969, pp. 187–88\)](#) considers it a CF, whereas for [Quirk et al. \(1985, pp. 1541–42\)](#), it is considered a prefix. Likewise, when *immuno-* is used, [Bauer \(2001, p. 70\)](#) states that it would be a compound or affixation in the case of *immuno-electrophoresis*, an affixation in the case of *immunodeficiency*, and a neoclassical compound in the case of *immunocyte*.

Furthermore, [Sánchez Fajardo \(2022, pp. 121–22\)](#) claims that both *-itis* and *-rrhea* “show a similar semantic reassignment that departs from strictly medical (or bodily) terms and results in abstract processes or tendencies (e.g., *lie-arrhea*, *opinionitis*) or mental states (e.g., *negorrhoea*, *queenitis*)”. Despite this semantic similarity, he regards *-itis* as a suffix and *-rrhea* as a CF and does not explicitly explain why they are differentiated in their morphological statuses.

2.3. Interim Summary

As mentioned above, neoclassical compounds have features not found in ordinary compounds, such as the occurrence of CFs and LVs. However, previous studies have either failed to account for them or have only assumed special operations for neoclassical compounds.

The issues concerning neoclassical compounding can be summarized in the following two questions:

- (10) a. What kind of word-formation rules are used to create neoclassical compounds?
- b. What is the morphological status of CFs that constitute neoclassical compounds?

The author’s answers to the above questions can be briefly stated as follows:

- (11) a. Neoclassical compounds are “constructions” in the form X-*o*-Y.
- b. CF is a term for the bound form used in the construction X-*o*-Y.

In (11a), X and Y are arbitrary elements, and *-o-* is a realization of LV. The “construction” refers to the construction in [Goldberg \(1995\)](#) and in research in CxM as developed by [Booij \(2010\)](#) and others.

Prior studies agree that CF, whether regarded as a lexeme or a special word-formation element, synthetically produces neoclassical compounds. However, the occurrence of LV and other specific properties of neoclassical compounds should be explained. For that purpose, CxM, which links formal, semantic, and phonetic information, is promising. Therefore, this paper proposes that neoclassical compounds are defined in the construction of X-*o*-Y, that neoclassical compounds consist of CFs, and that the specific properties of neoclassical compounds follow the properties of the construction. The tools used in the explanation are not assumed to be specific to the neoclassical compounds.

We also propose that the CFs that constitute the neoclassical compounds do not have an independent morphological status but are elements defined in this construction. In other words, the neoclassical compounds are “non-compounds”, because they are not formed

from lexemes, and “non-synthetic”, because they are not synthesized by combining CFs. In Section 4, we will show that viewing neoclassical compounds as such constructions can provide an explanation for their specificity. Before that, LV, a major formal feature of neoclassical compounds, will be examined in more detail in the next section.

3. The Status of Linking Vowels

As demonstrated in the many examples we have seen thus far, LVs are usually realized as *-o-*. There are four possibilities for treating this *-o-*, as discussed in Bauer (1998, p. 406):

- (12) a. phot-o-graph
 b. photo-graph
 c. phot-ograph
 d. photo-ograph → photograph

As Bauer (1998, p. 406) says, *-o-* “is viewed as a linking element between *phot* and *graph*”, as shown in (12a); it “is viewed as part of the first element”, as shown in (12b); it “is viewed as part of the second element”, as shown in (12c); or it “belongs to both the initial and final elements (as shown in [b] and [c]), and the sequence of *-oo-* is morphologically simplified to a single *-o-*”, as shown in (12d). Bauer himself states that “while (a) is the point of view usually taken by lexicographers, (b) appears to commend itself to native intuitions, in the sense that clippings invariably keep the *-o-*, for instance (*photo*)”. Prčić (2005, p. 318) supports this view, while Bauer and Huddleston (2002, p. 1662) essentially take the approach shown in (12d).

As another property, LV occurs when the FCF begins with a consonant, as in *gastrology*, but not when the FCF begins with a vowel, as in *gastritis*. This is shown in (13) as follows:

- (13) a. gastrology (<gastr-o-logy)
 b. gastritis (<gastr-itis)

Then, the occurrence of LV seems to be motivated by the phonological necessity to facilitate pronunciation.

In addition, there are cases where the LV is *-i-*, as shown in (14):

- (14) cone + fer → conifer
 insect + cide → insecticide
 herb + vore → herbivore (Scalise 1984, p. 76)

However, we argue that the LV in the neoclassical compound is specified as *-o-* in the construction, i.e., not any of (12), for the reasons discussed in the following subsections.

3.1. The Function of *-o-*

We mentioned the possibility that LVs are inserted for phonological necessity. However, when a neoclassical compound consists of three or more CFs, the LV is required even when the second CF begins with a vowel, as shown in the examples in (15). (The LV under consideration is double underlined in (15).)

- (15) a. gastroenteralgia, gastroenterology
 b. osteoarthritis

The assumption that the LV constitutes a single unit with the preceding CF is supported by the fact that there is a contraction ending in *-o-*, as exemplified in (16) (cf. Koga 2020):

- (16) photo (<photograph), cardio (<cardiovascular), chemo (<chemotherapy)

Bauer and Huddleston (2002, p. 1662) also state that in neoclassical compounds where the FCF is FM, the LV is omitted in relatively old words but is retained in more recent ones. Examples of the former are listed in (17), and those of the latter are in (18). Furthermore, there are some cases in which the LV is optional, like those shown in (19) (Bauer and

Huddleston 2002, p. 1662, fn 29). The year of the first occurrence of each word shown in the *OED* (2023) is indicated in parentheses. (In the subsequent examples, the number in parentheses after the word also indicates the year of its first occurrence shown in the *OED*.)

- (17) hom-organic (1854), micr-acoustic (1684), palaeontology (1833),
quadr-angle (a1398)¹
- (18) aero-engine (1907), autoerotic (1898), hydro-electric (1827),
microanalysis (1856), neo-impressionism (1892),
pseudo-intellectual (1880), psychoanalysis (1906)
- (19) ne(o)arctic (1858), palae(o)ichthyology (1872)

Although the number of words provided is small, the general trend is that words in (17), (19), and (18) began to be used in that order. In other words, when the FCF begins with a vowel, there seemed to be a phase in which the LV was omitted, followed by a phase in which the occurrence of the LV was optional, and then a phase in which the LV is retained. Then, the LV does not seem to occur from the phonological necessity but occurs to make a cohesive unit with ICF, which is followed by FCF.

3.2. Transition from *-i-* to *-o-*

LV in the neoclassical compound not only uses *-o-* but also *-i-*, as shown in (14) when the FCF is of Latin origin.² (14) is repeated in (20) as follows:

- (20) cone + fer → conifer
insect + cide → insecticide
herb + vore → herbivore (Scalise 1984, p. 76)

However, as shown in (21), *-o-* may be used even when the FCF is of Latin origins, such as *-cide* and *-meter*.

- (21) biometer; autocide, biocide, genocide

In addition, both *-o-* and *-i-* may be used interchangeably.

- (22) bactericidal (1878), bactericide (1884), bactericidally (1899), bacteriocidal (1943)

The form *bacterio-* is newer than *bacteri-*. In (22), the wavy-lined *bactericidal* and *bacteriocidal* show that the former originally had the vowel *-i-* in the linking part, and then, the latter appeared with *-o-*.

Examples of the uses of both *-o-* and *-i-* as LVs can also be found in the cases of *agro-* and *agri-*, as follows:

- (23) a. agribusiness (1955), agrobusiness (1958)
- b. agribusinessman (1961), agrobusinessman (1960)
- c. agrochemical (1953), agrochemical (1920)
- d. agri-industrial (1919), agroindustrial (1906)
- e. agri-industry (1950), agro-industry (1907)
- f. agriologist (1875-97), agrologist (1907)
- g. agriology (1878), agrology (1849)
- h. agri-tourism (1978), agro-tourism (1987)

Above, (23) enumerates the cases in the *OED* where both *agri-* and *agro-* are used as ICF with the same FCF. In (23), it is not obvious which is newer between *agri-* and *agro-*. However, when we look at the compounds listed in the *OED* that contain both *agri-* and *agro-*, it is words with *agro-* for which we find various examples. The examples in (24) and (25) are the complete list in the *OED* of compounds containing *agri-* and *agro-*, excluding derivatives.

- (24) agricultor (?a1425), agriculture (?1440), agricolation (1623), agrimotor (1916), agriproduct (1948), agriscience (1958), agri-food (1968), agri-monetary (1975), agrivoltaic (2011)
- (25) agronome (1776), agrogeology (1909), agrometeorology (1925), agrobiologie (1930), agrochimie (1930), agroecologie (1930), agrotechnologie (1932), agroforestry (1934), agrobacterium (1942), agro-ecosystem (1949), agro-city (1950), agro-town (1950), agrogorod (1951), agrodespotic (1957), agromechanization (1959), agro-politics (1960), agroliterate (1976), agroscape (1980), agroterrorism (1994)

Thus, the *agro-* form is more productive, and we can say that *agri-* in *agriculture*, which is one of the most common examples of the use of *-i-*, can be considered to be replaced by *agro-*. Also, considering the cases of *bactericidal* and *bacteriocidal* above, it seems that the productive form of LV in the neoclassical compound is *-o-*.

3.3. LV as a Constant in the Schema within CxM

In the next section, we propose that the productive form of neoclassical compounding is the X-o-Y form, in which *-o-* occurs regardless of the origin of the co-occurring FCF and not because of phonetic requirements. Regarding the cases in which *-i-* is used as LV, we claim that they are designated lexically and that the more general form of X-o-Y is preferred in forming new words.

4. Neoclassical Compounds in Construction Morphology

4.1. Specific Properties of Neoclassical Compounds

At least five characteristics of neoclassical compounds distinguish them from ordinary compounds.

- (26) a. They are composed of CFs.
 b. LV appears in them.
 c. Even in languages in which the left constituent is the head of the ordinary compounds, the right constituent is the head in the case of neoclassical compounds.
 d. The situations in which they are used are often restricted to scientific or academic ones.
 e. The position of the primary stress in the whole word is determined by the FCF.

Regarding Statement (26a), the status of CFs has been discussed in previous studies, as seen in Section 2. For the LV in (26b), we proposed that *-o-* is considered a productive form of the LV used in neoclassical compounds in Section 3.

As for the head mentioned in (26c), let us look at the examples provided by Amiot and Dugas (2020, p. 861). In French, the ordinary compound *homme-grenouille* consists of *homme* for “man” and *grenouille* for “frog”, and the meaning of the whole word is “frogman”, which is a kind of person. Then, the head of the whole compound is the noun *homme*, which occurs on the left-hand side of the word. In contrast, the neoclassical compound *phytothérapie* consists of *phyto-* for “plant” and *thérapie* for “therapy”. Since the overall meaning of this neoclassical compound is “phytotherapy”, which is a kind of therapy, the head is *thérapie*, and it occurs on the right-hand side. Thus, the position of the head differs in the ordinary compound and the neoclassical compound.

As for (26d), most examples we have seen thus far are in the scientific or academic field.

Concerning the position of the primary stress stated in (26e), the neoclassical compound differs from the ordinary compound in that the last element, FCF, determines the position of the primary stress in the neoclassical compound. In the case of ordinary compounds, the position of the primary stress may differ depending on the meaning, even in the same sequence of words. (The bold letter indicates that the primary stress is placed on that vowel.)

- (27) a. [government [[pay review] policy]]
 ‘government policy for reviewing pay’
 b. [[government pay] [review policy]]
 ‘policy for reviewing government pay’ (Spencer 1991, pp. 319–20)

Above, (27a) and (27b) show the noun phrase, which is composed of the same four words. When the sequence means “government policy for reviewing pay (=27a)”, the primary stress is placed on *pay*. However, the primary stress is placed on *review* when the sequence means “policy for reviewing government pay (=27b)”.³ In this respect, the neoclassical compound differs significantly. In the case of neoclassical compounds, the position of the primary stress is determined by the FCF. There are three types. The first type is the case in which the primary stress is placed on the FCF itself: *-itis* is this type, and the primary stress is placed on *-itis* itself regardless of the number of the preceding CFs. So, both *gastritis*, which has one preceding CF, and *gastroenterology*, which has two preceding CFs, have the primary stress placed on *-itis*. This is shown in (28), with the bold letter showing the primary stress.

- (28) *gastritis*, *gastroenteritis*

The second type is the *-logy* type, in which the primary stress is placed on the vowel immediately preceding the last CF, i.e., LV, and the third is the *-scope* type, in which the primary stress is placed on the CF immediately preceding the last CF. Below, (29) and (30) exemplify the *-logy* type and *-scope* type, respectively.

- (29) *gastrology*, *gastroenterology*, *biology*, *astrobiology*
 (30) *gastroscope*, *gastroenteroscope*

The examples in (29) show that in neoclassical compounds with the *-logy* type, the position of the primary stress does not change even if the number of CFs increases. This is true for *astrobiology*, which has potentially two different meanings (“biology in space” (<astro- + biology) and “the study of things living in space” (<astrobio- + -logy)). In both meanings, the position of the primary stress remains the same; in the case of the *-logy* type, the primary stress is placed on the LV.

On the other hand, in the *-scope* type, the primary stress is placed on the CF before the last CF, so the position of the primary stress changes as the number of CFs increases. In (30), in the case of *gastroscope* (<gastro- + -scope), the primary stress is placed on the *gastro-*, and in the case of *gastroenteroscope* (<gastro- + entero- + -scope), it is placed on the *entero-*.

Thus, in neoclassical compounds, the position of the primary stress of the whole word is determined by the type of the last CF, i.e., FCF, and unlike ordinary compounds, the semantic relations among ICFs and FCF are not related to their stress pattern.

4.2. Neoclassical Compounds in Construction Morphology

To explain the properties stated in (26), we propose that neoclassical compounds have the basic schema of (31) in the framework of Construction Morphology (CxM) (e.g., Booij 2010; Booij and Hüning 2014; Arcodia and Basciano 2018; Booij and Audring 2018).

- (31) <[X_i -o- Y_j]_{vk} ↔ [SEM_j with relation R to SEM_i]_k>
 [+bound] [+bound]

Condition: used in the scientific field

In (31), the part enclosed by < > is the schema, which indicates that the “form” on the left side of the bidirectional arrow corresponds to the “meaning” on the right side. The correspondence between the “form” and the “meaning” is indicated by the subscripts. The elements with the same subscripts correspond to each other. Also, (31) indicates that the two constituents, X and Y, are both bound forms because they bear the feature [+bound], that there is LV, -o-, between X and Y, that the head of the neoclassical compound, both in

Assuming that a combination of (32a-c) is derived through the change in features, one more form, i.e., FM-*o*-FM, is expected to occur.

(32) d. FM-*o*-FM

Indeed, this form does occur: it corresponds to what Prčić (2005) calls “quasi-neoclassical compositions”. The examples are shown in (37) as follows:

(37) semantico-pragmatic, lexico-grammatical, palato-alveolar, convexo-concave,
Euro-American, Czecho-Slovak, Serbo-Croatian (Prčić 2005, p. 321)

However, unlike the cases of (33) and (34), the right-hand elements in (37) are not the head of the whole words, which are non-endocentric. An example of an FM-*o*-FM form in which the right-hand element is the head is shown in (38) as follows:

(38) resistojet (Adams 1973, p. 130)

Above, (38) is an example in Adams (1973). The first part of *resistojet* corresponds to the verb *resist* (or to a part of the noun *resistor*), the second part to the noun *jet*, and the word as a whole refers to “an experimental electrothermal engine” (See Adams 1973, p. 130).

When both the first and second parts are FM, the structure is the same as that of an ordinary compound, except for the presence of LV. Thus, the presence of the LV means that the whole word is considered a neoclassical compound. Examples like (38) are not found often; in fact, this is the only example of a right-headed neoclassical compound in the form of FM-*o*-FM that I could find in the academic literature. However, it does exist, and this type might be motivated by non-endocentric expressions, such as those in (37).

Furthermore, Bauer and Huddleston (2002) and Baeskow (2004) have observed that, among subtypes of neoclassical compounds containing FM, those of the *bioscience* type (CF-*o*-FM) are more productive than those of the *hamburgerology* type (FM-*o*-CF). One would expect the schema for the former (the *bioscience* type), i.e., (33), to have been established earlier than the schema for the latter (the *hamburgerology* type), i.e., (34), and Hayashi (2023), which examined the examples listed in the *OED*, reported results that did not contradict this expectation.

Thus, the CxM approach is an effective and promising framework that can capture the specific properties of neoclassical compounds summarized in Section 4.1 and explain the relationships among their subtypes.

CxM has another advantage. As noted in the last part of Section 2.2, it is difficult to distinguish clearly between affixes and CFs. In CxM, both are treated as constants without distinction in the constructional idioms.

4.3. Remarks on Ordinary Compounds

Finally, a brief discussion of the formation of ordinary compounds will be in order. The schema in (31) is based on (39), which was proposed by Booij (2013) as “the most abstract schema for English compounds”.

(39) $[X_i Y_j]_{Y_k} \leftrightarrow [SEM_i \text{ with relation } R \text{ to } SEM_i]_k$ (Booij 2013, p. 258)

In schema (39), it is explicitly shown that this structure is both formally and semantically right-headed. In the formal structure shown to the left of the arrow, “the category of the word as a whole is identical to that of the right constituent, since they share the category variable *Y*”. In the semantic structure shown to the right of the arrow, the semantic head is co-indexed with the formal head. The example Booij gives us is *windmill*. His explanation is as follows:

- (40) A *windmill*, for instance, is a type of mill, not a type of wind, and it denotes a mill which has some semantic relation with the left constituent *wind*, namely ‘powered by.’
(Booij 2013, pp. 258–59)

Thus, Booij sees ordinary compounds as having the right-headed structure by means of a schema.

However, in the case of ordinary compounds, it is the free morphemes that occur in place of the variables X and Y in (39). Then, unlike in the case of neoclassical compounds, the schema for ordinary compounds does not need the feature [+bound] to indicate that the constituent is a bound form. It also does not require “the condition part” concerning the domain in which it is used, as seen in (31), (33), and (34). Furthermore, the mechanism of “extension through the feature alternations at the level of constructional idioms” that this paper argues for in neoclassical compound formation is irrelevant because there is no specification of the feature from the beginning. Thus, except for being able to explain right-headedness, there is no advantage in explaining ordinary compounds with a schema such as (39). In this respect, there is a significant difference between compounding and neoclassical compound formation.

Compounds in languages other than English should also be mentioned. The use of schemas and constructional idioms makes it clear that compounds have the right-headed structure, both formally and semantically, even in languages that have different morphological properties. For (39), it was a schema proposed by Booij for Germanic languages such as English, German, and Dutch. Similar schemas are also proposed by Booij and Hüning (2014) for expressions containing bound forms in German and Dutch, and by Arcodia (2011) and Arcodia and Basciano (2018) for those in Chinese. Thanks to this kind of research in CxM, it has become easier to discuss the similarities of compounds and the expressions containing bound forms in different types of languages. In the words of Arcodia and Basciano (2018, p. 249), “This means that cross-linguistic comparison may be carried out on the basis of actual constructions, rather than on categories which, often, may not be applied in a consistent way to typologically distant languages”.

From another perspective, free morphemes are listed as they are in the lexicon. In contrast, bound morphemes cannot occur alone, so they should be listed with additional conditions. In CxM, constructional idioms are used to specify the conditions under which they occur. In this way, CxM has the advantage over other frameworks of being able to explain bound forms without making any special assumptions.

One more point that should not be overlooked is that ordinary compounds in languages other than English also have vowels that function in a similar way to the LV discussed in this paper. For example, as Ralli (2013, p. 17) shows, in Greek compounds, *-o-* occurs systematically between two constituents (e.g., *psar-o-tavérna* “fish tavern” (<psár(i) “fish” + *tavérna* “tavern”), *patat-o-keftés* “potato roll” (<patát(a) “potato” + *keftés* ‘roll’). The same is true for Brazilian Portuguese (e.g., *cervej-o-chato* “beer snob” (<beer-LE-snob), which is an example of Nóbrega (2020, p. 112)), and many other languages, as discussed in detail in relation to recursion in Mukai (2013).

However, the use of the linking vowel in English differs from that in other languages mentioned above in that it is restricted to a limited domain of use, i.e., the scientific field. Rather, the linking vowel in English can be compared to a genitive marker in genitive compounds. English has a genitive marker (i.e., ‘s), as shown in (41).⁴

- (41) children’s book, women’s magazine, Mother’s Day (Mukai 2013, p. 39)

This use of ‘s is similar to the LV discussed in this paper in that they are both semantically empty. Although we cannot enter into the analysis of genitive compounds, the genitive marker still differs from the LV in neoclassical compounds in that the domain of use is not limited to some field, such as the scientific. In this sense, the LV in neoclassical compounds has a specific property. In order to capture the use of the linking vowel around a specific domain of use, it is useful to deal with in the framework of construction grammar, which

links form, meaning, and domain of use. Therefore, I have proposed (31) as a construction for neoclassical compounds.

Furthermore, as noted in (26c), for words with this LV in the scientific domain, the right constituent is the head, even in languages such as French, where the left constituent is the head of ordinary compounds. Thus, there may be a cross-linguistic construction for words containing LV in the scientific field. However, it is beyond the scope of this paper.

5. Conclusions

In this paper, to capture the specific properties of neoclassical compounds, we used the CxM framework and argued that neoclassical compounds are constructions of the form “X-o-Y”, in which X and Y are basically bound morphemes, but free morphemes are also used. Combining forms, which have been discussed in previous studies, are the bound forms that occur in the positions of X or Y in the construction “X-o-Y”. The combining form does not have independent status in morphology but is defined to be the element used in the construction “X-o-Y”. Originally, the combining form should be of classical origin. In fact, it is difficult to determine whether it is of classical origin or not, and we focus on the fact that it is a bound form since it also covers newly used forms such as *immuno-* (as in *immunology*). This view makes it possible to capture the commonalities with splinters in blends (e.g., *edu-* and *-tainment* in *edutainment*), which are also bound forms. However, this will be a subject for future research.

Going back to the analysis of neoclassical compounds in CxM, we see the hierarchical relationship between “constructions”, “constructional idioms”, and “words”, with different levels of abstraction, and their interactions explain the diversity and productivity of neoclassical compounds. Thus, although some issues remain, such as how to explain the fact that the stress pattern of the neoclassical compounds as a whole is determined by the FCF, the research of neoclassical compounds within the framework of CxM is shown to be effective and promising.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Acknowledgments: I would like to thank the academic editors and anonymous reviewers for their valuable comments and insightful feedback, which greatly helped to clarify the arguments of this paper. Any remaining errors are the responsibility of the author.

Conflicts of Interest: The author declares no conflict of interest.

Notes

- ¹ “a” before a date denotes *ante* (“before”).
- ² There are also other vowels used as linking vowels. Prčić (2005, pp. 316–17) mentions, in addition to *-o-* and *-i-*, *-y-* in *tachymeter*, *-u-* in *acupuncture*, and *-a-* in *genealogy* and *vibraphone* are also found in the position of ICF endings, and it says, “All these ICF endings function as linking vowels between the bare left-hand input element and its right-hand companion element and provide a smooth and euphonic transition from the first element to the other”.
- ³ More basically, Giegerich (2009, p. 188) provides examples where the same two-word expression, such as *toy factory* or *woman doctor*, can have the stress placed on the preceding word or on the following word, and the meaning changes depending on the position of the stress. For example, in *toy factory*, when the stress is on the following word, it means “factory which is a toy”, while when the stress is on the preceding word, the semantic relationship between the two words becomes more complex. The latter is a more typical stress pattern for compound words.
- ⁴ Mukai (2013, p. 39) cites examples in (41) from Shimamura (1986) and Taylor (1996).

References

- Adams, Valerie. 1973. *An Introduction to Modern English Word-Formation (English Language)*. London: Longman.
- Adams, Valerie. 2001. *Complex Words in English*. Harlow: Pearson Education Limited.

- Amiot, Dany, and Edwige Dugas. 2020. Combining forms and affixoids in morphology. In *The Oxford Research Encyclopedia*. Edited by Rochelle Lieber. Oxford: Oxford University Press, pp. 855–73.
- Arcodia, Giorgio Francesco. 2011. A Construction Morphology account of derivation in Mandarin Chinese. *Morphology* 21: 89–130. [CrossRef]
- Arcodia, Giorgio Francesco, and Bianca Basciano. 2018. The construction morphology analysis of Chinese word formation. In *The Construction of Words: Advances in Construction Morphology*. Edited by Geert Booij. Cham: Springer, pp. 219–53.
- Baeskow, Heike. 2004. *Lexical Properties of Selected Non-Native Morphemes of English*. Tübingen: Gunter Narr Verlag.
- 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: 403–22. [CrossRef]
- Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie, and Rodney Huddleston. 2002. Lexical word-formation. In *The Cambridge Grammar of the English Language*. Edited by Rodney Huddleston and Geoffrey K. Pullum. Cambridge: Cambridge University Press, pp. 1621–721.
- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Booij, Geert. 2013. Morphology in Construction Grammar. In *The Oxford Handbook of Construction Grammar*. Edited by Thomas Hoffmann and Graeme Trousdale. Oxford: Oxford University Press, pp. 255–73.
- Booij, Geert, and Jenny Audring. 2018. Partial motivation, multiple motivation: The role of output schemas in morphology. In *The Construction of Words: Advances in Construction Morphology*. Edited by Geert Booij. Cham: Springer, pp. 59–80.
- Booij, Geert, and Matthias Hüning. 2014. Affixoids and constructional idioms. In *Constructions All the Way Everywhere: The Extending Scope of Construction Grammar*. Edited by Ronny Boogaart, Timothy Coleman and Gijbert Rutten. Berlin: De Gruyter, pp. 77–103.
- Giegerich, Heinz. 2009. Compounding and Lexicalization. In *The Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 178–200.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
- Hayashi, Hiromi. 2023. Shin-Koten-Fukugogo ni Mirareru “Kobun” no Kakucho to Seisan-sei (Extension and Productivity of “Constructions” in Neoclassical Compounds). Ph.D. dissertation, University of Tsukuba, Tsukuba, Japan.
- Kastovsky, Dieter. 2009a. English word-formation, combining forms and neo-classical compounds: A reassessment. In *Current Issues in Unity and Diversity of Languages: Collection of the Papers Selected from the CIL*. Tokyo: LSK, Vol. 18, pp. 724–34.
- Kastovsky, Dieter. 2009b. Astronaut, astrology, astrophysics: About combining forms, classical compounds and affixoids. In *Selected Proceedings of the 2008 Symposium on New Approaches in English Historical Lexis (HEL-LEX 2)*. Somerville: Cascadilla Proceedings Project, pp. 1–13.
- Koga, Kentaro. 2020. *Goi to Bunpo no Aida de: Furansugo Fukugo-meishi no Seisan-teki-na Keisei-ho (Between the Lexicon and the Grammar: The Productive Formation of Compound Nouns in French)*. Tokyo: Surugadai-Shuppansha.
- Marchand, Hans. 1969. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*, 2nd ed. München: Verlag C. H. Beck.
- Mukai, Makiko. 2013. Recursive compounds and linking morpheme. *International Journal of English Linguistics* 3: 36–49. [CrossRef]
- Nagano, Akiko, and Masaharu Shimada. 2014. Morphological theory and orthography: Kanji as a representation of lexemes. *Journal of Linguistics* 50: 323–64. [CrossRef]
- Nóbrega, Vitor. 2020. No escape from categorization: An insider’s view of compounds. *Ilha Desterro* 73: 103–26. [CrossRef]
- OED. 2023. Oxford English Dictionary. Available online: <https://www.oed.com/> (accessed on 1 December 2023).
- Prčić, Tvrtko. 2005. Prefixes vs. initial combining forms in English. *International Journal of Lexicography* 18: 313–34.
- Prčić, Tvrtko. 2008. Suffixes vs. final combining forms in English: A lexicographic perspective. *International Journal of Lexicography* 21: 1–22.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Ralli, Angela. 2013. *Compounding in Modern Greek*. Dordrecht: Springer.
- Sánchez Fajardo, José Antonio. 2022. *Pejorative Suffixes and Combining Forms in English*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Scalise, Sergio. 1984. *Generative Morphology*. Dordrecht: Foris.
- Shimada, Masaharu, and Akiko Nagano. 2011a. A note on affixal and non-affixal anaphors in Japanese. In *Linguistic Theories and Their Application (WICLIS-2011)*. Seoul: Hankookmunhwasa.
- Shimada, Masaharu, and Akiko Nagano. 2011b. Zi-nouns in Japanese and related issues. *Studies in Language and Literature. Language* 59: 75–106.
- Shimamura, Reiko. 1986. Lexicalization of syntactic phrases. *English Linguistics* 3: 20–37. [CrossRef]
- Spencer, Andrew. 1991. *Morphological Theory*. London: Blackwell.
- Taylor, John R. 1996. *Possessives in English: An Exploration in Cognitive Grammar*. Oxford: Clarendon Press.

-
- Warren, Beatrice. 1990. The importance of combining forms. In *Contemporary Morphology*. Edited by Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer and John R. Rennison. Berlin: Mouton de Gruyter, pp. 111–32.
- Yoshioka, Ryo. 2011. Combining Forms in English and Their Morphological Status. Master's thesis, University of Tsukuba, Tsukuba, Japan.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.