

The Role of Combining Forms in Creating New English Compounds: Data-Driven and Construction Approaches*

Jin-hong Huang, *College of Foreign Languages,
Fujian Normal University, Fuzhou, China*
(jhhuang@fjnu.edu.cn) (<https://orcid.org/0000-0003-0233-789X>)
and

Yongwei Gao, *College of Foreign Languages and Literature,
Fudan University, Shanghai, China*
(Corresponding Author, ywgao@fudan.edu.cn)
(<https://orcid.org/0000-0001-7026-9273>)

Abstract: This paper explores the role of combining forms (CFs) in the formation of new compounds in contemporary English and examines their productivity and underlying mechanisms. This study applies the data-driven and construction approaches to investigate 11 neoclassical and native CFs, utilising both dictionaries and corpora when employing the data-driven approach. This paper culls new compounds from the *Oxford English Dictionary* and three neologism dictionaries. The three main findings are as follows: Firstly, the 11 CFs used in forming new compounds can be classified as highly productive, moderately productive, or low in productivity. Secondly, the construction morphology (CxM) can be used to analyse the formational mechanisms of new compounds. According to CxM, the CF compounds are abstracted as hierarchical schemas that are form–meaning pairs. The compounds exhibit syntactic and semantic constraints on their formation. Thirdly, the status of neologisms as compounds, blends or derivatives is scrutinised, along with their treatment in dictionaries. This study provides insights into the ongoing evolution of compounding in present-day English and discusses the role of CFs in lexical innovation.

Keywords: COMBINING FORMS, NEW ENGLISH COMPOUNDS, DATA-DRIVEN APPROACH, CONSTRUCTION MORPHOLOGY, PRODUCTIVITY, FORMATIONAL MECHANISM

Opsomming: Die rol van kombinasievorms in die skep van nuwe Engelse samestellings: Datagedrewe en konstruksiebenaderings. In hierdie artikel word die rol van kombinasievorms (KV's) in die vorming van nuwe samestellings in moderne Engels verken, asook hul produktiwiteit en onderliggende meganismes. Hierdie studie pas die datagedrewe en konstruksiebenadering toe om 11 neoklassieke en inheemse KV's te ondersoek, en sowel woordeboeke

* A version of this paper was presented at the 6th Globalex Workshop on Lexicography and Neology (GWLN-6), held on 3 July 2024 at the University of Pretoria, Hatfield Campus, Pretoria, South Africa.

as korpora word tydens die toepassing van die datagedrewe benadering gebruik. Nuwe samestellings uit die *Oxford English Dictionary* en drie neologismewoordeboeke word in hierdie artikel geselekteer. Die drie hoofbevindings is soos volg: Eerstens kan die 11 KV's wat gebruik word om nuwe samestellings te vorm, geklassifiseer word as hoogs produktief, matig produktief of laag in produktiwiteit. Tweedens kan die konstruksiemorfologie (KxM) gebruik word om die vormingsmeganismes van nuwe samestellings te ontleed. Volgens die KxM word die KV-samestellings geabstraheer as hiërargiese skemas wat uit vorm-betekenispare bestaan. Die samestellings toon sintaktiese en semantiese beperkings op hul vorming aan. Derdens word die status van neologismes as samestellings, reduksiesamestellings of afleidings ondersoek, sowel as hul hantering in woordeboeke. Hierdie studie verskaf insigte in die deurlopende evolusie van samestelling in hedendaagse Engels en bespreek die rol van KV's in leksikale innovasie.

Sleutelwoorde: KOMBINASIEVORMS, NUWE ENGELSE SAMESTELLINGS, DATAGEDREWE BENADERING, KONSTRUKSIEMORFOLOGIE, PRODUKTIWITEIT, VORMINGSMEGANISME

1. Introduction

Combining forms (CFs), a lexical concept that first appeared in the English language as early as 1884 and was defined by the *Oxford Dictionary of English* (ODE) as "a form of a word normally used in compounds in combination with another element to form a word", have traditionally played an important role in forming new words in the English language. Researchers have drawn a line between CFs and other morphological elements such as affixes, splinters, affixoids, and roots (Fradin 2000). Here it is necessary to distinguish CFs briefly from affixes and splinters. A CF can be used to attach to an independent word (e.g., autopilot), an affix (e.g., autocyst), or another CF (e.g., autography), while an affix can combine either a free word (e.g. self-made) or a CF (e.g., contranym)¹, but not another affix (e.g., *pre- + -ism). In terms of semantics, a CF tends to convey more specific meanings than an affix. The number of CFs has always been on the rise whereas there are seldom new affixes emerging in English. Splinters, the clipped elements of blends, are originally meaningless and non-independent (Adam 1973). Some splinters evolve into CFs through frequent use (e.g. Franken- from *Frankenstein*), however only those splinters that show high productivity and widespread usage tend to acquire this status. In this study, a CF is defined as follows: As a bound morpheme with an independent morphological status and an integral form and meaning, a CF can appear in compounds and blends and can be categorised as a 'neoclassical CF' or 'native CF' (cf. Section 2.1).

English compounds are words formed by combining two or more bases to create new lexemes with a distinct meaning. They can be written as a single word (e.g., microbreak), with a hyphen (e.g., micro-expression), or as separate words (e.g., micro scale). In present-day English, compounds are continually evolving as new combinations to reflect technological advancements, social changes, and cultural trends.

Traditionally, the productivity of most CFs pales in comparison with that

of affixes. However, a number of CFs are fairly productive in English today. CFs like *immuno-* and *-plasty* have been used to spawn at least one hundred scientific terms respectively. *Cyber-*, a relatively new CF, has been very productive in forming new English compounds whose number may have amounted to dozens or even over a hundred, as are recorded by major English dictionaries such as the ODE and *Collins English Dictionary* (CED). Frequently used *cyber-* compounds include *cyberbullying*, *cybercafé* and *cybercrime*, to mention a few. Other newly emerged CFs like *e-* (e.g., *e-book*) and *-preneur* (e.g., *dadpreneur*) have so far formed scores of new words that have been widely used in English. Thanks to their vitality, the current number of CFs has apparently far surpassed that of affixes, as is shown in the *Oxford English Dictionary* (OED) in which there are over two thousand CFs and only several hundred affixes, including both prefixes and suffixes.

A significant number of new compounds are created with CFs and other word-forming elements. While scholars such as Fischer (1998) and Mattiello (2023) have examined the productivity of CFs, the productivity of CFs in forming new compounds remains unexplored. Although a profusion of studies into compounds has been conducted from different perspectives (e.g., Scalise and Vogel 2010; Bauer 2017), little is known about the formational mechanisms of new English compounds formed with CFs. Given the current situation, this paper aims to examine these aspects. Specifically, this paper deals with 11 CFs (see Table 1 in Section 4.1). The two research questions are as follows: (1) What is the productivity of the 11 CFs in forming new compounds? (2) What is the mechanism underlying the formation of new compounds formed with the CFs?

A data-driven approach in linguistic studies emphasises the analysis of real-world language use through large datasets, such as dictionaries and corpora, which facilitate the identification of empirical evidence for linguistic phenomena. To address these research questions, this paper employs a data-driven approach involving dictionaries and corpora (see Section 4.3) to examine both the productivity of CFs in forming new compounds and the formational mechanisms of such compounds. The construction morphology proposed by Booij (2010) is applied to analyse the formational mechanisms of these CFs.

2. Previous studies on CFs and English compounds

2.1 Research on CFs

CFs are traditionally defined as neoclassical elements of Greek or Latin origin. Scholars have explored whether neoclassical CFs should be classified as instances of affixation or compounding (Bauer et al. 2013: 431-490). Some researchers extend the definition of CFs to include elements originating from native English words (i.e., 'native CFs' in Wiemeyer's (2019) terms), whereas Amiot and Dugas (2020) divide these word-forming elements into classical CFs and modern CFs. CFs also can be grouped by their positions in words, mainly including ini-

tial combining forms (ICFs) and final combining forms (FCFs) (Plag 2003: 156). Approaches to classifying CFs therefore vary (see Mattiello 2023).

Apart from the classifications of CFs, prior studies focus on their semantics (e.g., Lehrer 1998; Lalić-Krstin et al. 2022) and productivity (e.g., Fischer 1998; Rita-Kasari 2013). While these studies have greatly enhanced the understanding of CF classification, semantics, and productivity, they devoted less attention to how CFs function semantically and productively when forming compounds. This study aims to address this research gap by investigating the formation, semantic properties, and productivity of new CF compounds.

2.2 Research on English compounds

Previous studies into English compounds have concentrated on their classifications, word forms or structures, semantics, word classes, as well as their formational mechanisms. Scholars have been rather inconsistent in their classification of compounds, not only in terms of the classes deemed important, but also in the nomenclature used for the classes (e.g., Fábregas and Scalise 2012). While noun compounds have been extensively studied, systematic empirical research on adjectival and verbal compounds remains scarce (e.g., Bauer et al. 2013). Structural analyses have shed light on the internal organisation of compounds, considering aspects such as headedness, constituency, stress, and orthography (e.g., Anderson 1992). In addition, prior studies have distinguished between idiomatic and compositional compounds and have applied various theoretical frameworks to analyse their meanings (e.g., Ten Hacken 2016). Research on word classes has provided detailed categorisations of compound nouns, compound verbs, compound adjectives, and even less commonly discussed categories such as compound adverbs and compound interjections (e.g., Adams 1973). Scholars have also examined the mechanisms underlying the formation of compounds, employing perspectives from schema theory, blending theory, or paradigmatic analysis (e.g., Fauconnier and Turner 2002; Bagasheva 2020).

Despite the valuable insights mentioned above, studies specifically examining CF compounds across these dimensions are limited (e.g., Díaz Negrillo 2014; Pulcini 2020). In particular, new compounds involving CFs are an underexplored area. Booij (2009: 201-216) applies construction morphology to analyse compounds in different languages, partly involving English compounds. However, only a few studies have adopted this approach to exploring neoclassical compounds (cf. Hayashi 2024). This paper therefore uses the approach of construction morphology to examine the compounds formed with CFs.

3. Theoretical considerations

3.1 Construction Morphology in word formation

Construction Morphology (CxM), a morphological theory developed by the Dutch

linguist Geert Booij, integrates insights from Construction Grammar and introduces the concept of 'construction' into morphological analysis. The notion of 'construction' in CxM is derived from Goldberg's (2006: 5) hypothesis, which characterises constructions as linguistic patterns whose form or function is not strictly predictable from their constituent parts. The basic ideas of CxM have been explicated and defended in Booij (2010), and in a number of introductory chapters on CxM (e.g., Booij 2018). In CxM, a word is conceptualised as a pairing of form and meaning, with words in the lexicon understood as systematic form–meaning correspondences (Booij 2010: 5).

3.1.1 Word-based morphology

A distinctive feature of CxM is its word-based approach to morphology, which contrasts with the morpheme-based approach (Kiparsky 1982). While the morpheme-based approach sees complex words as concatenations of morphemes, word-based morphology treats the word itself as the model for new words. In other words, CxM posits that "words are formed from words" (Aronoff 1976: 46). Within the framework of generative word formation, Aronoff (1976) introduces the theory of "word-based morphology", which hypothesises that "all regular word formation processes are word-based. A new word is formed by applying a regular rule to a single already existing word" (ibid: 21). CxM considers each word as a holistic entity and examines its internal structure by comparing it with paradigmatically related words. There is no strict separation of grammar and lexicon (Booij 2018: 3). For instance, instead of viewing *writer* as a simple combination of write and *-er*, CxM sees it as part of a broader family of words (e.g., *singer* and *teacher*) that English speakers recognise and use productively.

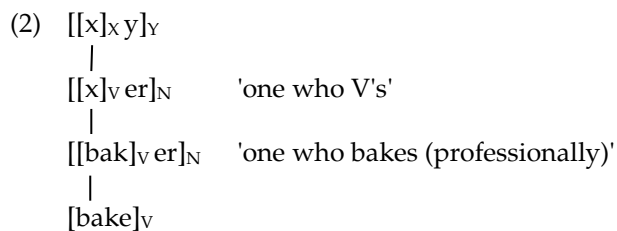
3.1.2 Schemas

Unified schemas that allow for the derivation of multiple complex word formation products are a central concept in CxM. The model of CxM uses constructional schemas to account for the systematic form–meaning relations between words. Booij (2015: 450) asserts that a "schema expresses the systematic and abstract form–meaning correspondence found in a set of word pairs". Schemas are used to represent the internal structures of complex words. The relationship between a constructional schema and its instantiations can be illustrated through an inheritance tree, with more abstract schemas occupying the higher nodes, and specific lexical items situated at the lower nodes. The lexical item inherits properties from the higher-level schema. An English suffixal schema (1), as outlined by Booij (2018: 4), can serve as an example to demonstrate this relationship:

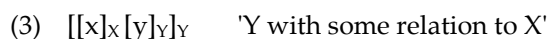
- (1) $[[x]_{Ni} y]_{Aj} \leftrightarrow [\text{possessing characteristic properties of SEM}_i]_{SEM_j}$

The double arrow in the schema (1) symbolises the correspondence between form and meaning. The variable x denotes the phonological form of the noun. Co-indexation indicates that the semantic representation (SEM) of the noun is embedded within the meaning of the corresponding adjective. This schema assumes an independently specified semantic representation of the noun, thus highlighting that the schema itself is grounded in paradigmatic relations among lexical items. The variables i and j stand for the lexical indexes on the phonological and syntactic properties of words.

Hierarchical constructional schemas represent generalisations about existing complex words while simultaneously providing a structured framework for the formation of new words (Booij 2010). The relationship between an abstract constructional schema and its individual instantiations can be represented in a tree structure, where the schema functions as the dominant node, and individual lexical items occupy lower-level nodes that inherit the properties of their superordinate categories. For instance, the derivative *baker* can be represented within such a hierarchical structure (Booij 2005: 124), as shown in schema (2):



Booij (2005: 123) uses schema (3) to represent a compound. For compounds formed with ICFs and FCFs, they can be abstracted as $[[\text{ICF}]_x [y]_Y]_Y$ and $[[x]_X [\text{FCF}]_Y]_Y$ respectively, where the form of the variable x or y is not entirely predictable. Therefore, both $[[\text{ICF}]_x [y]_Y]_Y$ and $[[x]_X [\text{FCF}]_Y]_Y$ can be identified as constructions in the framework of CxM. CxM regards compounds as constructions, where form–meaning correspondences are not strictly predictable. By abstracting CF compounds into schema (3), CxM offers a flexible and systematic way to examine the formation of new compounds. Moreover, it provides a clear method for understanding the formational mechanisms of these compounds, as the schemas allow for the variation of components while preserving the underlying structural and semantic patterns. Through the inheritance of properties across different levels of abstraction, CxM can effectively account for both the consistency and innovation observed in the formation of new compounds.



In comparison with the schema theory (Bybee 1985) or the paradigmatic approach to compounds (Bauer 2017), words in CxM are treated as form–meaning pairings, and CxM deals with the semantic relationships between elements. The schema

theory or the paradigmatic approach, on the other hand, focuses on the abstract structures and patterns that produce new words within established paradigms, with less emphasis placed on meaning as a fundamental factor in word formation.

3.1.3 Semantic and syntactic constraints on compounds

CxM asserts that constructional schemas impose semantic constraints on the formation of compounds, so that the meaning of the resulting compound remains interpretable within systematic lexical patterns. For example, the CF *photo-* in *phototherapy* is subject to the constraint that the second element must be semantically associated with treatment or the medical field. The application of semantic constraints in CxM follows prototypical patterns, which account for the varying productivity of CFs in forming new compounds. A case in point is *tele-*, which exhibits high compatibility with communication-related nouns (e.g., *telephone*) but does not typically co-occur with physical object nouns like *telechair*. On the other hand, CxM maintains that syntactic constraints imposed by constructional schemas ensure the syntactic predictability of compounds. In this regard, the CF *geo-* in *geography* is subject to the constraint that the second element can be a noun or a CF, thereby producing a noun compound. The systematic nature of syntactic constraints in CxM shows that CFs combine only bases that conform to specific word class patterns.

3.2 Morphological productivity

Morphological productivity, a key concept in the word-formation research and morphological theories, is the likelihood that a construction will apply to a new item (Bybee 2010: 94). Each lexical slot in a construction has its own degree of productivity. If a process in morphology is fully regular and actively used in the creation of new words, this process is considered productive (Richards and Schmidt 2010: 463). Fischer (1998: 63) argues that the current productivity of CFs can be measured by the number of new words included in dictionaries. Compilers of dictionaries largely consider the established words for inclusion. Bauer (2001: 159) takes the view that using dictionaries for the measurement of productivity is "likely to prevent relatively rare established words being taken as neologisms, and it allows the genuinely productive use of morphological processes to be seen".

Both qualitative and quantitative methods can be used to assess productivity. This paper adopts the quantitative measure and counts the number of new compounds included in dictionaries over specific periods to analyse the productivity of CFs in forming new compounds. Counting the number of new words in a given phase is a measure of productivity (Plag 2021: 489). However, the use of dictionaries is not immune from problems. As Bauer (2001: 158-159) notes, con-

sulting a dictionary provides a conservative idea about what is 'part of the language' by often excluding neologisms, rare words, and established words deemed invaluable to the target audience.

According to Booij (2010: 50), productivity is a type of holistic property of morphological constructions. Schemas in the context of morphological productivity apply to both productive and non-productive form–meaning correspondences and show that regularity can exist without productivity. However, a unified schema enhances the productivity of word-formation patterns by reinforcing morphological generalisations. Unified schemas provide a framework for understanding key differences in the evolution of individual subpatterns, particularly with regard to morphological productivity and the semantic dimensions of word-formation constructions (Kempf and Hartmann 2018: 442). In turn, productivity is linked to type frequency, which influences the degree of a schema's entrenchment (Barðdal 2008).

4. Methodology and procedures

4.1 Research objectives

The 11 CFs examined in this study and their origins are listed in Table 1. By referring to the OED, the dates of first occurrence of the said CFs in English formations were obtained. One exception of a CF not included in the OED but included in the list is *-ware*, derived from 'software'. Based on their etymological origins, the 11 CFs are classified into neoclassical CFs and native CFs, while their roles as ICFs or FCFs are also identified.

Table 1: CFs examined in this study

CF	Origin	Date of first occurrence in English formations	Type	
bio-	Latin	the early 19th century	ICF	neoclassical CF
macro-	Greek	the 1820s		
micro-	Greek	1603		
techno-	Greek	the 19th century		
tele-	Greek	the first half of the 19th century		
-phobia	Latin	the 18th century	FCF	neoclassical CF
-plasty	Greek, combined with the suffix <i>-y</i>	the first half of the 19th century		
cyber-	cybernetics	the 1960s	ICF	native CF
e-	electronic	the late 1980s		
immuno-	immune + <i>-o-</i>	the beginning of the 20th century		
-ware	software	the 1960s	FCF	

As allomorphic variants of classical Latin or Greek words, the five ICFs *bio-*, *macro-*, *micro-*, *techno-*, and *tele-* that end with a vowel /o/ or /i/ are neoclassical CFs. Two other neoclassical CFs *-phobia* and *-plasty* are FCFs. This paper also examines four native ICFs, including *cyber-*, *e-*, *immuno-*, and *-ware*, of which *immuno-* derives from 'immune' and *-o-* while the other four are truncated from existing English words. *Cyber-*, *e-*, and *-ware* are new CFs whereas the other eight CFs have been in use longer. The 11 CFs were selected because they provide a representative sample including both traditional and emerging CFs, diverse origins, structural roles (ICFs and FCFs), and distinct historical stages of development. All of the said CFs have been productive in making new words to varying degrees in the past decades. These CFs are well-suited for analysis, as they have formed a sufficient number of new compounds for investigation. The time span from the 1960s to the 2020s was chosen because sufficient time has passed to enable a measurement of productivity of the 11 CFs in forming new compounds.

4.2 A data-driven approach

This paper employs a data-driven approach, which means "methods based on statistical models learnt from data" (Mitkov 2022: 1252). Mittelberg et al. (2007: 30) emphasise that 'data-driven' does not equal 'data-based', since in data-driven work the data are not just the material basis for the analysis, or just used to test or exemplify theories formulated prior to the advent of large corpora. Rather, the data actually guide the course of analytical processes and theory building. Corpora have been particularly useful in enabling the creation of data-driven models (Baker et al. 2006: 41). Dictionaries compiled using large corpora of real-world texts exemplify a data-driven approach because they derive word meanings and usage patterns from empirical evidence, rather than theoretical or prescriptive rules. Both dictionaries and corpora are utilised when employing the data-driven approach in the present study.

4.3 Sources of data

The data sources of this paper are dictionaries and corpora. Firstly, the new compounds formed with the 11 CFs were selected from major English dictionaries, such as the OED. The OED is used because it is a comprehensive and authoritative resource on the English language, tracing the meanings, history, and evolution of words. Other online dictionaries do not systematically provide the dates of emergence of entries, which is essential for measuring the productivity of CFs in forming new compounds during certain periods. Therefore, only the OED was selected among large-size online dictionaries.

Secondly, this study makes good use of three English–Chinese dictionaries of neologisms to collect new compounds. *A Dictionary of Contemporary English Neologisms and Trendy Words* (DCEN, Gao 2024) records nearly 17 000 entries of

present-day English neologisms and trendy terms, most of which have emerged or gained widespread usage since the 1970s (1970s–2020s). This dictionary particularly focuses on recording new words that have been frequently used in the media over the past decade. *An English–Chinese Dictionary of Neologisms in the New Era* (ECDN, Gao 2023) includes over 4 000 new English words and phrases that have appeared or been frequently used since the beginning of the 21st century (2000–2015). *A 21st Century English–Chinese Dictionary of Neologisms* (CECDN, Gao 2021) documents more than 2 500 new words and meanings that have been active in use since the early 2000s (2000s–2020s). For this paper, the three dictionaries were chosen because of their thorough and timely recording of English neologisms across several decades, which allows for a detailed study of new compounds formed with CFs. The use of these four dictionaries to collect new compounds also has limitations. The OED's recording of new words may not be timely, and some emerging compounds can be missing. Similarly, the three English–Chinese neologism dictionaries may not cover all new compounds, despite their provision of comprehensive and representative samples.

Thirdly, this study utilises three corpora: the Corpus of Historical American English (COHA; Davies 2010), the Corpus of Contemporary American English (COCA; Davies 2008), and the News on the Web (NOW; Davies 2016). These corpora were used to determine the first recorded dates of new compounds. COHA, comprising over 475 million words, provides historical American English data from 1820 to 2019 that includes texts from novels, magazines, newspapers, and nonfiction. However, it lacks spoken data where the new compounds in question may appear. To address this, COCA was included, as it contains over one billion words from various genres, including spoken texts. NOW, a monitor corpus with 16,4 billion words from online newspapers and magazines (2010–present), updates monthly with 200–220 million words. As new compounds frequently emerge in news and media, NOW is particularly useful for tracking their development.

The OED is particularly valued for its ability to pinpoint the first recorded use of words and thus the dates of first appearance of compounds can be attested to make sure whether they are neologisms. The appearance date of new compounds enables us to explore the productivity of CFs during a certain span of time. If a compound is not included in the OED, the authors refer to other online dictionaries that provide the etymological information where may contain the dates of first use of headwords. The online dictionaries in question involve ODE, CED, *Merriam-Webster's Collegiate Dictionary* (MWCD), and *Wiktionary* (WKD). If the dates of first appearance are all absent in the five dictionaries (i.e. OED, ODE, CED, MWCD, and WKD), reference is also made to the three corpora (COHA, COCA, and NOW) to attest their dates of first use in the dataset.

4.4 Data compilation

Each of the CFs was searched in the OED. The OED records all the said word-

forming elements as CFs, except *-ware*. The OED provides the section of "compounds and derived words" for the CFs, and the types of formation are also indicated. In the case of *bio-*, its formations included in the OED are 357 compounds, 6 derivatives, and 4 borrowings. The numbers of compounds formed with the CFs covered in the OED are listed in Table 2. The numbers of compounds in the three neologism dictionaries are also investigated, as is shown below.

Table 2: Number of compounds in four dictionaries

	OED	DCEN	ECDN	CECDN
Time span of inclusion	1960–2024	1970–2024	2000–2015	2000–2021
CF	Compound number in each dictionary			
bio-	357	125	29	9
macro-	143	12	0	4
micro-	635	110	30	24
techno-	57	18	2	2
tele-	201	21	3	2
-phobia	93	30	11	1
-plasty	42	15	2	6
cyber-	68	104	23	6
e-	26	64	28	4
immuno-	83	16	0	0
-ware	13	34	17	1

5. Results and analysis

5.1 Dates of first use for new compounds

The new compounds formed with each CF and recorded in the OED, DCEN, ECDN, and CECDN were compiled together, with duplicates excluded. The dates of first use for these compounds are sourced from dictionaries such as the OED, ODE, CED, MWCD, and corpora like COHA, COCA, and NOW. Most of the CFs in question were used to form compounds before 1960 or have formed such words since the 1960s (except *e-*). This paper uses 1960 as the starting year to measure the productivity of the CFs in producing new compounds while the ending year is 2024. Compounds formed between 1960 and 2024 and recorded in the aforementioned four dictionaries are enumerated in Table 3. The '2020s' are used to refer to the period from 2000 to 2024, for consistency with others. With the data-driven periodisation in place, the analysis of these compounds are divided into ten-year intervals. Note that the absence of new CF compounds in certain periods in the four dictionaries does not imply that these CFs did not form any compounds during these periods. Instead, it indicates that the com-

pounds they formed did not become established to warrant lexicographic inclusion.

The *bio-* formations within English were found from the early 19th century. The OED, DCEN, ECDN, and CECDN include 214 *bio-* compounds that appeared before the 1960s. A surge occurred in the 1960s and 1970s, with 44 and 48 new compounds, respectively. Despite fluctuations, 168 new *bio-* compounds that have emerged since the 1960s are included in the aforementioned dictionaries. *Macro-* first appeared in English formations in the 1820s, and a total of 108 *macro-* compounds that emerged before the 1960s are listed in the four dictionaries. The 1960s saw 22 new *macro-* compounds included, followed by fewer than five being added per decade from the 1980s onward, totaling 42 since the 1960s. *Micro-* was first seen in use in English words as early as 1603. The four dictionaries have included as many as 508 *micro-* compounds that appeared before the 1960s. The 1960s saw the peak of new *micro-* compounds with 57, followed by a gradual decline. 176 new *micro-* compounds that emerged from the 1960s to the 2020s are recorded in the four dictionaries. The other eight CFs showed varying degrees of productivity, as indicated in Table 3.

Table 3: Number of new compounds in four dictionaries by dates of first appearance

CF	Undetermined	Before 1960s	1960s	1970s	1980s	1990s	2000s	2010s	2020s	Sum of 1960–2024
bio-	1	214	44	48	29	18	8	20	1	168
macro-	0	108	22	7	4	2	2	4	1	42
micro-	2	508	57	33	22	16	12	35	1	176
techno-	0	22	9	10	13	4	1	5	0	42
tele-	0	163	10	14	9	2	0	3	0	38
-phobia	0	78	4	1	3	5	2	9	1	25
-plasty	0	37	3	1	0	2	1	6	0	13
cyber-	3	0	4	4	7	51	9	27	2	104
e-	0	0	0	0	5	24	18	21	0	68
immuno-	0	37	29	14	3	0	0	0	0	46
-ware	1	0	1	2	9	5	4	7	3	31

5.2 The formal properties of new compounds

This section examines the formal properties of new compounds formed with 11 CFs from the 1960s to the 2020s and explores the distribution of solid, hyphenated, and open compounds among these compounds (see Table 4). With the exception of *techno-* and *e-*, the other nine CFs show a strong tendency towards forming solid compounds. The number of solid compounds formed with *bio-*, *micro-*, as well as *cyber-* surpasses 100, demonstrating the established roles of these CFs in such compounds, as in *bioabsorption*, *microbeer*, and *cyberaggression*. CFs such as *tele-*, *-phobia*, *-plasty*, *immuno-*, and *-ware* exclusively form solid compounds. The

large number of solid compounds reflects their dominance in fields such as biology, medicine, science, and technology.

Table 4: Forms of new compounds recorded in the four dictionaries

CF	Solid compound	Hyphenated compound	Open compound
bio-	164	4	0
macro-	35	7	0
micro-	161	15	0
techno-	19	22	0
tele-	38	0	0
-phobia	25	0	0
-plasty	13	0	0
cyber-	101	3	0
e-	1	67	0
immuno-	46	0	0
-ware	31	0	0

Hyphenated compounds are less common but remain significant for certain CFs. For example, *techno-* exhibits a higher proportion of hyphenated compounds than solid ones. This suggests that *techno-* compounds often represent newly emerging or less standardised terms, such as *techno-junkie* and *techno-thriller*. In a similar vein, *e-* displays a striking preference for hyphenation, with 67 hyphenated compounds like *e-economy* and *e-payment*. None of the CFs in the dataset form open compounds. Although compounds such as *cyber cheating* and *cyber school* are included in the dictionaries, 'cyber' here functions as an independent word instead of a CF. They were not taken into account in this discussion.

5.3 The word classes of new compounds

The data in Table 5 demonstrate that the majority of new CF compounds are nouns. CFs such as *bio-*, *micro-*, and *cyber-* are particularly productive in forming noun compounds. Examples include *biocapacity*, *microangiography*, and *cyber-security*. Adjective compounds (e.g., *immunoregulatory* and *microaerobic*), while less common, are still notable, with CFs like *bio-*, *micro-*, and *immuno-* forming over 15 adjectives. In contrast, verb compounds (e.g., *cyberstalk* and *micromanage*) are rare among CFs, with *bio-*, *micro-*, *tele-*, and *cyber-* contributing only a handful of examples. Only *immuno-electrophoretically*, an adverb compound of *immuno-*, is recorded. The new compounds of *macro-*, *techno-*, and *-plasty* recorded are nouns or adjectives while those of *-phobia*, *e-*, and *-ware* are nouns.

Table 5: Word-classes of new compounds formed with CFs

CF	noun	adjective	verb	adverb
bio-	134	27	7	0
macro-	32	10	0	0
micro-	154	18	4	0
techno-	30	11	0	0
tele-	34	1	3	0
-phobia	25	0	0	0
-plasty	12	1	0	0
cyber-	95	4	5	0
e-	68	0	0	0
immuno-	26	18	1	1
-ware	31	0	0	0

Table 6 illustrates the compounding of CFs with other constituents, with the "CF + noun" construction dominating across all ICFs. *Bio-* forms 134 noun compounds, while *micro-* and *cyber-* contribute 154 and 91, respectively. Though "CF + adjective" and "CF + verb" constructions appear, they are less frequent, and the "CF + adverb" construction occurs only once in *immunoelectrophoretically*. Some ICFs, such as *techno-*, *cyber-*, and *immuno-*, also form "CF + CF" constructions. As for the FCFs, *-phobia* mainly forms "noun + CF" constructions and the only *-phobia* compound in the form of "CF + CF" is *coronophobia*. The dominant construction of *-plasty* is "CF + CF". Compounds formed with *-ware* exclusively use the "noun + CF" construction.

Table 6: Constituents of new compounds formed with CFs

ICF	CF + noun	CF + adjective	CF + verb	CF + adverb	CF + CF
bio-	134	27	7	0	0
macro-	32	10	0	0	0
micro-	154	18	4	0	0
techno-	30	9	0	0	2
tele-	34	1	3	0	0
cyber-	91	4	5	0	4
e-	68	0	0	0	0
immuno-	25	18	1	1	1
FCF	CF + CF	adjective + CF	noun + CF	/	/
-phobia	1	0	24	/	/
-plasty	9	1	3	/	/
-ware	0	0	31	/	/

5.4 The semantic properties of new compounds

5.4.1 The semantic preference of CF compounds

ICFs and FCFs combine other elements to form lexemes. Such elements are termed the 'first element' (e.g., *intersex* in *intersexphobia*) and the 'second element' (*aerosol* in *bioaerosol*) respectively. Exploring the semantic categories of the first or second elements of the CF compounds can reflect their semantic preference. For example, Fischer (1998: 146) presents nine categories of the second elements of *cyber-*, including persons, living beings, objects, body parts or sense organs, places or buildings, and so on. The new compounds formed with the CFs from the 1960s to the 2020s are analysed below. This study classifies the first or second element into five semantic categories: objects, abstractions, behaviours and activities, people, and animals. The "animal" here denotes any living creature other than a human being. These five categories can cover all the first or second elements of the 11 CFs.

Table 7: Semantic types of first or second elements of CF compounds

CF	Objects		Abstractions		Behaviors & Activity		People		Animals	
	number	proportion	number	proportion	number	proportion	number	proportion	number	proportion
bio-	46	27.3%	72	42.8%	41	24.4%	9	5.3%	0	0
macro-	14	33.3%	24	57.1%	2	4.7%	1	2.3%	1	2.3%
micro-	66	37.5%	68	38.6%	32	18.1%	6	3.4%	4	2.2%
techno-	3	7.1%	27	64.2%	3	7.1%	9	21.4%	0	0
tele-	5	13.2%	13	34.2%	16	42.1%	4	10.5%	0	0
-phobia	5	20.0%	7	28.0%	4	16.0%	8	32.0%	1	4.0%
-plasty	10	76.9%	2	15.3%	1	7.6%	0	0	0	0
cyber-	5	4.8%	34	32.7%	37	35.5%	27	25.9%	1	0.9%
e-	29	42.6%	14	20.5%	19	27.9%	6	8.8%	0	0
immuno-	5	10.8%	24	52.1%	17	36.9%	0	0	0	0
-ware	1	3.2%	15	48.3%	15	48.3%	0	0	0	0

Bio- primarily combines words expressing abstractions (*bioavailability*), or those denoting objects (*biochip*), and behaviours and activities (*biohacking*). Only a few *bio-* compounds involve people (*biohacker*), and none refers to animals. *Macro-* and *micro-* display clear semantic preferences for abstractions (*macromodelling*) and objects (*macrolens*). *Micro-* also combines the second elements referring to behaviours and activities (*microblogging*). *Techno-* strongly favors the combination with the second elements representing abstractions (*technocentrism*). *Tele-* mainly combines with the second elements representing behaviours and activities (*telebroking*). *Cyber-* primarily combines the second elements that belong to behaviours and activities (*cyberauction*), and abstractions (*cybercommunity*). Further data on other CFs can be found in Table 7.

5.4.2 The semantic change of the CFs in compounds

The meanings of 11 CFs in Table 8 illustrate their dynamic evolution, as these CFs have adapted to new contexts and domains when forming new compounds. By examining the origins of CFs and their current usage, insight is gained into

how these morphemes reflect cultural, scientific, and technological developments. If *cyber-* is considered as an example, it is evident that *cyber-* means "computers" in early compounds found in the OED, like *cyberculture*. It appears in nine new compounds with this meaning, as in *cyberathlete*, *cybergeek*, and *cyberphobe*. Terms like *cyberart*, *cybersphere*, and *cyberbabe* reflect its use to describe information technology, focusing on digital systems or products. In 63 new compounds, *cyber-* conveys this meaning. Meanwhile, compounds such as *cyberspace* and *cyberlife* extend the meaning of *cyber-* to virtual reality and online existence. *Cyber-* in compounds such as *cybercrime* and *cyberwarfare* signifies the application of digital technologies to illegal or hostile activities. Furthermore, *cyberaggression*, *cyberstalking*, and *cyberscam* reflect the influence of digital and virtual environments on social activities.

Table 8: Meanings of CFs and the number of new compounds with this meaning

CF	Meaning	No.	Meaning	No.	Meaning	No.	Meaning	No.	Meaning	No.
bio-	living beings	146	biology	18	life	2	biochemical substances as weapons	2		
macro-	large-scale	28	large	12	large compared with the norm	1	visible	1		
micro-	small	157	microscopy	10	abnormally small	5	microphotography or microprinting	2	one millionth	2
techno-	of technology	28	craft, art	7	technological	7				
tele-	to or at a distance	30	done by means of the telephone	6	relating to television	2				
-phobia	fear	14	hostility, prejudice	7	aversion	4				
-plasty	formation	10	plastic surgery	3						
cyber-	information technology	63	virtual reality	32	computers	9				
e-	sth. in an electronic format	30	computer-mediated or internet-based	27	electronic financial transactions	11				
immuno-	immune	22	immunity	20	immunology	4				
-ware	software	28	software (in a metaphorical sense)	3						

6. Discussion

6.1 The productivity of CFs in creating new compounds

As demonstrated in Table 3 and Section 5.1, the productivity of the 11 CFs in forming new compounds varies significantly between neoclassical and native CFs from the 1960s to the 2020s. *Bio-*, *micro-*, and *cyber-*, each with over 100 new compounds, are highly productive. *Macro-*, *techno-*, *tele-*, *e-*, *immuno-*, and *-ware*, with more than 30 new formations, show moderate productivity, while *-phobia* and *-plasty*, with fewer than 30 new formations, exhibit low productivity. The thresholds of 30 and 100 new compounds here provide practical reference points rather than absolute indicators.

The rate at which the 11 CFs form new compounds also notably differs across decades. CFs like *bio-*, *micro-*, and *cyber-* exhibit a rapid expansion, especially

during periods of significant scientific and technological development. For instance, new *bio-* compounds experienced a sharp increase in the 1960s and 1970s, a period of intense biological research and medical advancements. These CFs quickly adapted and expanded as new concepts emerged, which indicates that the pace of forming compounds is closely tied to societal changes and technological revolutions.

6.2 The formational mechanism of CF compounds

In CxM, word formation is analysed through hierarchical constructional schemas that reflect systematic form–meaning correspondences. Using *cyber-* as an example, the unified schema $[[\text{cyber}]_X[\text{y}]_Y]_Y$ governs its overall combinatory potential, where *cyber-* combines another element *y* to form a lexeme that belongs to the word class *Y*. In accordance with the analysis in Section 5.4.2, *cyber-* has acquired three meanings so far. Therefore, the schema of *cyber-* can express the meaning of "something related to information technology, virtual reality, and computers". At a more specific level, this general schema branches into $[\text{cyber}-[\text{y}]_{\text{Ni}}]_{\text{Nj}}$ (94 instances, e.g., *cyberart*), $[\text{cyber}-[\text{y}]_{\text{Ai}}]_{\text{Aj}}$ (4 instances, e.g., *cybersecure*), $[\text{cyber}-[\text{y}]_{\text{Vi}}]_{\text{Vj}}$ (5 instances, e.g., *cyberloaf*), and $[\text{cyber}-[\text{y}]_{\text{Ci}}]_{\text{Nj}}$ (4 instances, e.g., *cybernaut*). Each of these schemas represents a distinct form–meaning pairing, with the lower-level schemas inheriting properties from the more abstract schema. This hierarchical structure demonstrates how CxM accounts for both consistency and variation in the formation of new CF compounds.

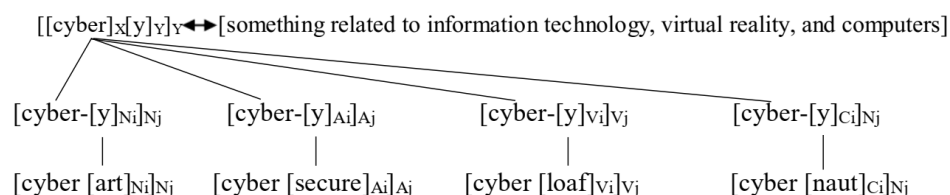


Figure 1: The constructional schema of *cyber-* compounds

Table 9 presents the constructional schemas of the other 10 CFs, and their schemas have varying degrees of productivity in the formation of new compounds. The CF compounds show syntactic constraints, since the first or the second elements which the CFs can combine are of different word classes. It can be noticed that *bio-*, *micro-*, and *tele-* all have three schemas. They can combine a noun, an adjective, or a verb to form new compounds. CFs such as *techno-* and *-plasty* also exhibit three schemas. Apart from compounding with a noun or an adjective, they can form compounds with CFs. *Macro-* can be abstracted into two schemas and this CF can form compounds with a noun or an adjective. These constraints reflect the interplay between form and meaning in CxM, where

some CFs serve highly specialised functions while others exhibit broader productivity across multiple schemas.

Table 9: The schematicity of CFs and the number of their new compounds

CF	Schema	No.	Schema	No.	Schema	No.	Schema	No.
cyber-	[cyber- y] _{Ni}] _{Nj}	94	[cyber- y] _{Ai}] _{Aj}	4	[cyber- y] _{Vi}] _{Vi}	5	[cyber- y] _{Ci}] _{Nj}	4
bio-	[bio- y] _{Ni}] _{Nj}	149	[bio- y] _{Ai}] _{Aj}	31	[bio- y] _{Vi}] _{Vi}	7	/	/
immuno-	[immuno- y] _{Ni}] _{Nj}	26	[immuno- y] _{Ai}] _{Aj}	18	[immuno- y] _{Vi}] _{Vi}	1	[immuno- y] _{ADi}] _{ADj}	1
	[immuno- y] _{Ci}] _{Nj}	1	/	/	/	/	/	/
macro-	[macro- y] _{Ni}] _{Nj}	33	[macro- y] _{Ai}] _{Aj}	10	/	/	/	/
	/	/	/	/	/	/	/	/
micro-	[micro- y] _{Ni}] _{Nj}	167	[micro- y] _{Ai}] _{Aj}	20	[micro- y] _{Vi}] _{Vi}	4	/	/
techno-	[techno- y] _{Ni}] _{Nj}	31	[techno- y] _{Ai}] _{Aj}	10	[techno- y] _{Ci}] _{Nj}	2	/	/
tele-	[tele- y] _{Ni}] _{Nj}	37	[tele- y] _{Ai}] _{Aj}	1	[tele- y] _{Vi}] _{Vi}	3	/	/
e-	[e- y] _{Ni}] _{Nj}	68	/	/	/	/	/	/
-phobia	[[x] _{Ci} -phobia] _{Nj}	35	/	/	/	/	/	/
-plasty	[[x] _{Ni} -plasty] _{Nj}	3	[[x] _{Ai} -plasty] _{Aj}	1	[[x] _{Ci} -plasty] _{Cj}	14	/	/
-ware	[[x] _{Ni} -ware] _{Nj}	33	/	/	/	/	/	/

As analysed in Section 5.4.1, CF compounds exhibit semantic constraints based on the semantic types of elements that they combine. *Macro-*, *micro-*, *cyber-*, and *-phobia* can combine the elements representing objects, abstractions, behaviours and activities, people, and animals. In contrast, the second elements of *bio-*, *techno-*, *tele-*, and *e-* exclude animals, whereas those of *immuno-* exclude both people and animals. Similarly, the first elements of *-plasty* and *-ware* do not fall within the categories of people or animals.

Highly productive schemas such as [bio- y]_{Ni}]_{Nj} and [micro- y]_{Ni}]_{Nj} illustrate how frequently used constructions become entrenched in the lexicon, while less productive schemas like [immuno- y]_{Vi}]_{Vi}, [tele- y]_{Ai}]_{Aj}, and [[x]_{Ai}-plasty]_{Aj} represent emerging or marginal patterns. The hierarchical organisation of schemas ensures that word-formation generalisations remain systematic, even as new compounds are created. The constraints observed in CF compounds highlight the importance of semantic and syntactic factors in determining a CF's combinatory potential.

6.3 Demarcation between compounds and others and their lexicographical treatment

In accordance with the data in the OED, DCEN, ECDN, as well as CECDN, CF formations involve compounds and blends. The following section examines the etymological treatment of CF formations in the OED, CED, ODE, MWCD, and WKD, as these five dictionaries provide such information for some entries.

The existing deficiencies are as follows: Firstly, the etymological information, especially the ways of formation provided for the same CF formation, may vary in the said dictionaries. Secondly, the etymological information offered for some CF formations is not sufficiently clear and accurate. For instance, *bioethicist* is a derivative. However, its way of formation is treated differently in the five above-mentioned dictionaries. *Bioethicist* is a compound according to the OED, whereas it is indicated as being derived from *bioethics* in ODE, MWCD, and WKD. As is

attested in the OED, *bioethics* first appeared in 1970, while the date of first appearance of *bioethicist* is 1973. It can be reasonably asserted that *bioethicist* is a derivative rather than a compound. Table 10 displays the lexicographical treatment of *cyberphobic* (compound), *biochar* (blend) and other words as well.

This study on CF compounds shows that CxM transcends rigid distinctions among compounds, derivatives, and blends by representing them through unified schemas, which reduces ambiguity in morphological classification. However, as is displayed in Table 10, dictionaries adopt traditional classifications and present word-formation information based on structural analysis. Given that CF formations can arise from different morphological processes, it is essential for dictionary compilers to propose a standardised lexicographical annotation for CF compounds, distinguishing them clearly from derivatives and blends to ensure the authority and accuracy of dictionaries. For CF compounds like *cyberphobic*, which is formed by two CFs, the etymology section should explicitly label them as CF compounds, along with the analysis of word-forming elements (see Table 10). Entries such as *bioethicist*, which involve the addition of a conventional affix to a pre-existing compound or root, should be treated under derivation. Dictionary compilers should specify the base (*bioethics*) and the attached affix (*-ist*), clarifying that the process is affixal rather than compositional. Regarding blends such as *biochar*, they should be labeled as blends, with the source words clearly identified. Since word-formation mechanisms can be complex, it is desirable for dictionaries to avoid overgeneralisation and acknowledge uncertainties where necessary, while incorporating traditional classification alongside insights from CxM.

Table 10: The etymological treatment of CF formations in dictionaries

CF formation	Suggested treatment	OED	CED	ODE	MWCD	WKD
cyberphobic	CF compound, from <i>cyber-</i> (comb. form) + <i>-phobic</i> (comb. form)	by derivation < <i>cyberphobia</i> n.	derivative of <i>cyberphobia</i>	/	/	from <i>cyber-</i> + <i>-phobic</i>
bioethicist	derivative, from <i>bioethics</i> + <i>-ist</i>	compounding, <i>bio-</i> comb. form + <i>ethicist</i> n., after <i>bioethics</i>	/	derivative of <i>bioethics</i>	derivative of <i>bioethics</i>	from <i>bioethics</i> + <i>-ist</i> , analogous to <i>ethicist</i>
microlender	derivative, from <i>microlending</i> + <i>-er</i>	compounding, <i>micro-</i> comb. form + <i>lender</i> n. compare <i>microlending</i> n.	/	derivative of <i>microlending</i>	derivative of <i>microlending</i>	from <i>micro-</i> + <i>lender</i>
biochar	blend, from <i>biomass</i> and <i>charcoal</i>	compounding, <i>bio-</i> (in <i>biomass</i> n.) + <i>char-</i> (in <i>charcoal</i> n.)	/	blend of <i>biomass</i> and <i>charcoal</i>	/	from <i>bio-</i> + <i>char</i>
cyberchondriac	blend, from <i>cyber-</i> (comb. form) + <i>hypochondriac</i>	compounding, <i>cyber-</i> comb. form + <i>-chondriac</i> (in <i>hypochondriac</i> adj.)	derivative of <i>cyberchondria</i>	blend of <i>cyber-</i> and <i>hypochondriac</i>	/	blend of <i>cyber-</i> and <i>hypochondriac</i>
e-tailer	blend, from <i>e-</i> (comb. form) + <i>retailer</i>	compounding, <i>e-</i> comb. form + <i>-tailer</i> (in <i>retailer</i> n.) compare <i>e-tailing</i> n.	/	blend of <i>e-</i> and <i>retailer</i>	derivative of <i>e-tail</i>	blend of <i>e-</i> + <i>retailer</i>

7. Conclusion

Based on data from four dictionaries — the OED, DCEN, ECDN, and CECDN — the findings reveal the productivity of the 11 CFs in forming new compounds from the 1960s to the 2020s. The rate at which the CFs create new compounds notably differs across decades: *bio-*, *micro-*, and *cyber-* are highly productive; *macro-*, *techno-*, *tele-*, *e-*, *immuno-*, and *-ware* are moderately productive; and *-phobia* and *-plasty* are low in productivity.

It is found that the CF compounds, governed by both syntactic and semantic constraints, can be abstracted as hierarchical schemas to represent form–meaning correspondences. In terms of syntactic constraints, the combinatory patterns of CFs vary, with some able to attach to multiple word classes and others restricted to specific word classes. The semantic constraints of CFs differ, with some combining broadly across categories like objects, abstractions, behaviours and activities, people, and animals, while others show selective restrictions, excluding certain categories, such as people or animals. This study also finds that the 11 CFs have undergone semantic changes and have acquired established meanings.

This study contributes to the understanding of compounds formed with CFs. Unlike previous studies focusing on general compounds, newly coined CF compounds are examined, shedding light on ongoing lexical innovation. By integrating the data-driven and construction approaches, this research provides insights into the realised productivity of the CFs in forming new compounds and their underlying mechanisms.

Acknowledgements

We would like to express our sincere gratitude to the two anonymous reviewers and the journal editors for their valuable suggestions and insightful comments.

Endnote

1. *Contra-* is a prefix, and *-onym* is a CF.

References

Dictionaries and tools

Collins English Dictionary Online (CED).

<https://www.collinsdictionary.com/> [14 March 2025]

Davies, M. 2008. The Corpus of Contemporary American English (COCA).

<https://www.english-corpora.org/coca/> [14 March 2025]

Davies, M. 2010. The Corpus of Historical American English (COHA).

<https://www.english-corpora.org/coha/> [14 March 2025]

- Davies, M.** 2016. *The News on the Web (NOW)*.
<https://www.english-corpora.org/now/> [14 March 2025]
- Gao, Y.W.** 2021. *A 21st Century English–Chinese Dictionary of Neologisms (CECDN)*. Shanghai: Fudan University Press.
- Gao, Y.W.** 2023. *An English–Chinese Dictionary of Neologisms in the New Era (ECDN)*. Beijing: Commercial Press.
- Gao, Y.W.** 2024. *A Dictionary of Contemporary English Neologisms and Trendy Words (DCEN)*. Beijing: Foreign Language Teaching and Research Press.
- Merriam-Webster's Collegiate Dictionary Online (MWCD)**.
<https://www.merriam-webster.com/> [14 March 2025]
- Oxford Dictionary of English Online (ODE)**.
<https://premium.oxforddictionaries.com/english/> [14 March 2025]
- Oxford English Dictionary Online (OED)**.
<https://www.oed.com/> [14 March 2025]
- Wiktionary (WKD)**.
<https://en.wiktionary.org/> [14 March 2025]

Other literature

- Adams, V.** 1973. *An Introduction to Modern English Word-formation*. London: Longman.
- Amiot, D. and E. Dugas.** 2020. Combining Forms and Affixoids in Morphology. Lieber, R. (Ed.). 2020. *The Oxford Encyclopedia of Morphology*: 855-873. Oxford: Oxford University Press.
- Anderson, S.R.** 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, M.** 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Bagasheva, A.** 2020. Paradigmaticity in Compounding. Fernández-Domínguez, J., A. Bagasheva and C. Lara-Clares (Eds.). 2020. *Paradigmatic Relations in Word Formation*: 21-48. Leiden/Boston: Brill.
- Baker, P., A. Hardie and T. McEnery.** 2006. *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barðdal, J.** 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.
- Bauer, L.** 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, L.** 2017. *Compounds and Compounding*. Cambridge: Cambridge University Press.
- Bauer, L., R. Lieber and I. Plag.** 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Booij, G.** 2005. Compounding and Derivation: Evidence for Construction Morphology. Dressler, W.U., F. Rainer, D. Kastovsky, and O. Pfeiffer (Eds.). 2005. *Morphology and Its Demarcations*: 109-132. Amsterdam/Philadelphia: John Benjamins.
- Booij, G.** 2009. Compounding and Construction Morphology. Lieber, R. and P. Štekauer (Eds.). 2009. *The Oxford Handbook of Compounding*: 201-216. Oxford/New York: Oxford University Press.
- Booij, G.** 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Booij, G.** 2015. Word Formation in Construction Grammar. Müller, P.O., I. Ohnheiser, S. Olsen and F. Rainer (Eds.). 2015. *Word Formation: An International Handbook of the Languages of Europe*. Vol. 1: 188-202. Berlin/New York: De Gruyter.

- Booij, G. (Ed.).** 2018. *The Construction of Words: Advances in Construction Morphology*. Cham: Springer.
- Bybee, J.** 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, J.** 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Díaz Negrillo, A.** 2014. Neoclassical Compounds and Final Combining Forms in English. *Linguistik Online* 68(6): 3-20.
- Fábregas, A. and S. Scalise.** 2012. *Morphology: From Data to Theories*. Edinburgh: Edinburgh University Press.
- Fauconnier, G. and M. Turner.** 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Fischer, R.** 1998. *Lexical Change in Present-day English: A Corpus-based Study of the Motivation, Institutionalisation, and Productivity of Creative Neologisms*. Tübingen: Gunter Narr Verlag.
- Fradin, B.** 2000. Combining Forms, Blends and Related Phenomena. Doleschal, U. and A.M. Thornton (Eds.). 2000. *Extragrammatical and Marginal Morphology*: 11-59. München: Lincom Europa.
- Goldberg, A.E.** 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Hayashi, H.** 2024. A Construction Morphology Approach to Neoclassical Compounds and the Function of the Linking Vowel. *Languages* 9(4):1-15.
- Kempf, L. and S. Hartmann.** 2018. Schema Unification and Morphological Productivity: A Diachronic Perspective. Booij, G. (Ed.). 2018. *The Construction of Words: Advances in Construction Morphology*: 441-474. Cham: Springer.
- Kiparsky, P.** 1982. From Cyclic Phonology to Lexical Phonology. Hulst, H.V.D. and N. Smith (Eds.). 1982. *The Structure of Phonological Representations. Part 1*: 131-175. Dordrecht: Foris.
- Lalić-Krstin, G., N. Silaški and T. Đurović.** 2022. Meanings of *-nomics* in English: From *Nixonomics* to *Coronanomics*, How *-nomics* Has Extended Its Original Meaning to Additional Senses. *English Today*: 1-8.
- Lehrer, A.** 1998. Scapes, Holics, and Thons: The Semantics of English Combining Forms. *American Speech* 73(1): 3-28.
- Mattiello, E.** 2023. *Transitional Morphology: Combining Forms in Modern English*. Cambridge: Cambridge University Press.
- Mitkov, R. (Ed.).** 2022. *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press.
- Mittelberg, I., T.A. Farmer, and L.R. Waugh.** 2007. They Actually Said That? An Introduction to Working with Usage Data through Discourse and Corpus Analysis. Gonzalez-Marquez, M., I. Mittelberg, S. Coulson and M.J. Spivey (Eds.). 2007. *Methods in Cognitive Linguistics*: 19-52. Amsterdam/Philadelphia: John Benjamins.
- Plag, I.** 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
- Plag, I.** 2021. Productivity. Aarts, B., A. McMahon and L. Hinrichs (Eds.). 2021. *The Handbook of English Linguistics*. Second edition: 483-499. Hoboken, NJ, USA: Wiley-Blackwell.
- Pulcini, V.** 2020. The Combining Form *multi-* in English Compounds. Maci, S., M. Sala and C. Spinzi (Eds.). 2020. *Communicating English in Specialised Domains. Festschrift for Maurizio Gotti*: 301-315. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Richards, J.C. and R.W. Schmidt.** 2010. *Longman Dictionary of Language Teaching and Applied Linguistics*. Fourth edition. London: Pearson Education.

- Rita-Kasari, E.** 2013. *The Morphological Productivity of Selected Combining Forms in English*. Unpublished MA Thesis. Helsinki: University of Helsinki.
- Scalise, S. and I. Vogel (Eds.).** 2010. *Cross-Disciplinary Issues in Compounding*. Amsterdam: Benjamins.
- Ten Hacken, P. (Ed.).** 2016. *The Semantics of Compounding*. Cambridge: Cambridge University Press.
- Wiemeyer, L.** 2019. The Diachronic Productivity of Native Combining Forms in American English. Wiegand, V. and M. Mahlberg (Eds.). 2019. *Corpus Linguistics, Context and Culture*: 223-252. Berlin/Boston: Walter de Gruyter.