The Development of ESP Lexicon Through New Combining Forms

Elisa Mattiello[1]

[1] Department of Philology, Literature and Linguistics, University of Pisa, Pisa, Italy

Correspondence: Elisa Mattiello, Department of Philology, Literature and Linguistics, University of Pisa, Via Santa Maria 67, 56126 Pisa, Italy. E-mail: elisa.mattiello@unipi.it

Abstract

This paper investigates the role of new combining forms in the formation of neologisms which are currently expanding the lexicon of English for Special Purposes (ESP). In the past, only neoclassical combining forms, such as initial *bio-* or final *-logy* (in *biology*), were productively used in ESP. Nowadays, specialized combining forms also include abbreviated forms of existing words (e.g., *cyber-* from *cybernetic* in *cyber-attack*), as well as secreted (i.e. reinterpreted) forms (e.g., *-bot* from *robot* denoting 'a type of automated program or software' in *knowbot*). The paper explores a set of combining forms attested since the second half of last century in the online version of the *Oxford English Dictionary* (OED) with the aim to demonstrate how specialized sectors, such as science or information technology, are being enriched by series of combining-form combinations. The paper conducts quantitative analyses in the *Corpus of Contemporary American English* (COCA) and the *News on the Web Corpus* (NOW) to substantiate the frequency and stability of specialized combining forms and their profitability in the formation of both novel and nonce words.

Keywords: neoclassical/abbreviated/secreted combining forms, ESP, word-formation

1. Introduction

Because of new scientific discoveries, medical advancements, and technological inventions, the lexicon of English is in continuous expansion (Arndt-Lappe et al., 2018). Many new words are recognized and repeatedly used by the speech community, therefore becoming part of English vocabulary as accepted neologisms. Others are coined for specific occasions and do not always become institutionalized or lexicalized (Brinton & Closs Traugott, 2005), yet their creation suggests how language may be enriched with new temporary vocabulary, namely nonce terms, connected to a historical, socio-political or economic period. The lexicon of English for Special Purposes (ESP), however, is rarely ephemeral, since the emergence of new specialized terms is generally motivated by lexical gaps which need to be filled.

The present paper analyses 30 combining forms retrieved from the *Oxford English Dictionary* (OED) and contributing to the lexical innovation of ESP. Combining forms (henceforth, CFs) are traditionally defined as neoclassical elements of Greek or Latin origin, such as *bio-* (Gr. *βίο*, Lat. *bio* 'life') or *-logy* (Gr. *λογία*, Lat. *logia* 'science'), which are added to other elements, either initially, as in *biodata* 'one's curriculum vitae', or finally, as in *musicology* 'the branch of knowledge that deals with music'. However, as recently demonstrated by Mattiello (2022), CFs comprise not only neoclassical bound elements, but also abbreviated forms, such as *e-* (from *electronic*) in *e-book*, *e-journal*, and secreted ones, such as *-nomics* (from *economics*, denoting 'the economic policies of a President or head of state') in *Clintonomics*, *Obamanomics* (cf. "classical" vs. "modern combining forms" in Prćić, 2005, 2008; Amiot & Dugas, 2021).

While in the literature neoclassical CFs are generally regarded as cases of compounding (Bauer, 2017, p. 150), abbreviated and secreted CFs have been attached different labels, including "affixoids" or "pseudo-affixes", which rather stress their derivational nature. Scholars also tend to confuse or merge CFs with the notion of "splinters" used in the formation of blends (Bauer, Beliaeva, & Tarasova, 2019, p. 62), thus confirming their "fuzzy" nature and difficult categorization. In a recent study (Mattiello, 2022), the three categories of neoclassical, abbreviated, and secreted CFs have been considered as part of transitional morphology, i.e. intermediate between two subcomponents of word-formation: i.e., derivation and compounding. Like affixes, CFs are bound morphemes added to bases, yet like compound constituents, they have a high lexical density and can be stressed (e.g., *e-* has the same meaning as the adjective *electronic* from which it is clipped and is stressed

in *é-mail*).

The dividing line between one category and the other is often not so well defined, nor are CFs classified in a univocal way in the linguistic literature. Section 2 tries to clear up the terminological confusion about CFs and related concepts. Section 3 explains the methodology followed for the selection and analysis of the data. Section 4 offers a qualitative scrutiny of the selected CFs and a quantitative analysis of the combining-form combinations that they produce, discriminating between highly frequent neological formations and occasionalistic nonce words. Finally, section 5 summarizes the research conducted and draws up some conclusions.

2. Combining Forms: Definitions and Terminological Distinctions

Warren (1990, p. 112) was the first to stress that "combining forms are morphemes of a rather special kind", which represent a quite heterogeneous group of elements, differing in their origin as (1) allomorphs of model words (e.g., *astro-* in *astrodome*), (2) truncated forms of model words (e.g., *cyber-* in *cyberphobia*), or (3) parts of model words which happen to coincide with existing words (e.g., *-gate*, from *Watergate*, denoting 'an actual or alleged scandal' in *sexgate*). The same CF can even differ in the way it is interpreted depending on the word to which it is added: e.g., in *cheeseburger* 'hamburger topped with cheese', *-burger* is an abbreviated form, whereas in *fishburger* 'fried patty made of fish served in a bread-bun', it is secreted, in that a *fishburger* does not contain any beef meat. Moreover, CFs are divided into initial and final (*e-* vs. *-nomics*) depending on their position either to the left or to the right of a base. With regard to position, there are very few exceptions (i.e. *morph-/-morph* and *phil-/-phile*), which can occur both in initial (e.g., *morph-ology*, *phil-ology*) and final positions (e.g., *anthropo-morph*, *anglo-phile*) (Bauer, 1983; Plag, 2003). Initial CFs are close to prefixes while final CFs are close to suffixes.

This heterogeneity manifests itself in terminological confusion (e.g. "affixoids", "pseudo-affixes", "semi-affixes", "affix-like formatives", "folkmorphs", "splinters", along with "combining forms") and descriptive vagueness. Bauer already in the 1980s claimed that CFs belong to a type of word-formation which "has received very scant attention in the literature on morphology" (Bauer, 1983, p. 213). Scholars indeed tend to relegate this topic to the category of neoclassical CFs and only include discussions on neoclassical compounds in their description of English word-formation, disregarding completely the modern categories. More recently, Kastovsky (2009, p. 12) has even argued that the notion of CF is not necessary at all, as "[t]he categories of 'word', 'stem', 'affix', 'affixoid', 'clipping' and 'blending' necessary in word-formation for independent reasons are sufficient to deal with the formations in question".

Following recent studies on the topic (e.g., Amiot & Dugas, 2021), Mattiello (2022) has dealt with this phenomenon showing the importance of a finer-grained classification and an in-depth description of CFs. From her analysis, CFs consist of a miscellaneous set of morphological elements which share features such as semantic and phonological weight with compound constituents, but also features such as boundedness and fixed position (either before or after a base) with affixes. For these reasons, Mattiello (2022) accommodates the different categories of CFs under the same umbrella term of "transitional morphology", straddling the demarcation line between two subcomponents of word-formation, i.e. derivation and compounding.

In line with Mattiello (2022), CFs are here defined as initial or final bound morphemes which are either allomorphic variants of classical Latin or Greek words (e.g., *bio-* above), or shortenings of (native or non-native) English words (e.g., *econo-* from *economic* in *econo-politics*), often with the intervention of a secretion process (e.g., *-aholic* 'person addicted to' in *work-aholic* 'person addicted to work'). These are respectively called "neoclassical", "abbreviated", and "secreted CFs", and the complex words obtained from them are called "combining-form combinations".

A notion which is often confused with CFs is that of affixoids. The term "affixoid" is used within Construction Morphology by Booij and Hüning (2014), who note how the element *free* has developed the more general meaning 'without what is denoted by the base word' when it is used as the right constituent of compounds, as in *sugar-free* and *fat-free*. Moreover, unlike the suffix *-less* having the same denotative meaning in *sugar-less* and *fat-less*, the affixoid *-free* also presupposes a slightly negative evaluation (e.g., that the presence of sugar is not good for one's health). Therefore, affixoids exist alongside formally identical and usually free 'parent' morphs; yet, they acquire a more generalized meaning (Stevens, 2005, p. 73) or undergo a "desemanticization" process (Amiot & Dugas, 2021), as *-free* 'clear of something which or someone who is regarded as problematic' in *trouble-free*, *pollution-free*, *risk-free*, etc. Thus, like CFs, affixoids are "productive" (Bauer, 2001), i.e., are part of schematic constructions with high type frequency (e.g., [N-*free*]), and are bound to their bases (e.g., in *sugar-free*, the suffixoid and the base cannot be split up as **sugar very free*). However, unlike CFs, affixoids

correspond to free-standing words, whereas in the case of CFs, correspondence to a free morpheme is accidental or synchronically irrelevant (cf. the CF -*gate* 'political scandal' vs. the noun *gate*).

Still another crucial distinction is between CFs and blend splinters. Splinters are word parts that merge with one another or with full words in a blend (e.g., in *jeggings ← jeans + leggings* there are two splinters which graphically overlap, while in *Brexit ← British + exit* only the first element is shortened). Hence, the notion of splinter is very close to that of CF, differing from it only in their diverse degrees of productivity (Bauer, 2001). We may even envisage a diachronic evolution from blend splinter to abbreviated or secreted CF. Many splinters are indeed transitional from their status as blend constituents (i.e. word parts) to a more productive affix-like status as CFs (Correia Saavedra, 2013, 2016). This evolution is manifest in -*aholic*, which was originally attested in blends (e.g., *work-aholic* [1947], *choco-holic* [1961]), then became a productive CF included in the OED with the meaning 'a person addicted to –' (e.g., *carbo-holic* [1973], *shop-aholic* [1977], *news-aholic* [1979], *spend-aholic* [1982]), with a higher degree of frequency, productivity, and stability.

In this paper, the focus is on the CFs which have become productive and stable in ESP, because of their reiterated use in the formation of novel words by experts in specialized areas.

3. Dataset and Methodology

The methodology adopted for the selection of the CFs examined in this study combined an advanced search in the OED, using the designated tool, with a manual cleaning of irrelevant examples. The advanced search was conducted in November 2023. First, I selected the tab 'entries' of the OED and entered the search text 'combining form'. This search returned a list of 2,297 matching entries, which were then put in chronological order, by arranging results by 'date (newest first)'. Since my focus was primarily on recent forms, I added a temporal parameter to my search, by selecting the range 1950-today, and concentrated on this dataset. The 77 results obtained were the starting point for my analysis. To this list *robo-* was added manually because, even if its earliest attestation is before 1950, in 1988 it became a secreted form referring to something or someone 'resembling a robot, esp. in being resilient, emotionless, or futuristic'. Another CF that deserved inclusion, but did not match my advanced search because it is not dated in the OED, is -*nomics*. I decided to incorporate it because the earliest formation displaying this CF is dated 1969.

The final 79 results include all three categories – namely, neoclassical, abbreviated, and secreted forms – yet with some elements having evolved in time from abbreviated to secreted. This dataset was then manually reduced to 30 – 18 initial and 12 final – specialized CFs for the purposes of my analysis. Among the selected CFs, some are apparently similar in form, but actually diverge in meaning: e.g., *e-* is abbreviated from either *electric* (e.g., in *e-car*) or *electronic* (e.g., in *e-journal*). Hence it counts as two separate CFs (*e-¹* vs. *e-²*) used in different specialized fields. The specialized fields of the selected CFs mainly involve digital and information technology, electronics, economics, and less relevantly chemistry, medicine, physiology, palӨontology, and linguistics.

In the corpus-based analysis, the three categories of neoclassical, abbreviated, and secreted CFs will be treated separately, also distinguishing between initial and final for each type, in order to verify their frequency and profitability in terms of new words obtained through this word-formation mechanism. Both type and token frequencies of occurrence in the *Corpus of Contemporary American English* (COCA) (Note 1) and in the *News on the Web Corpus* (NOW) (Note 2) will be shown, as well as their meanings and contexts of use. By observing quantitative data, a distinction between stable neologisms and nonce formations or hapax legomena (i.e. once-only attestations) will be also made. By exploring qualitative results, an evolution towards abbreviated and secreted CFs will be observed in ESP, as already remarked in general English (Mattiello, 2022).

4. Specialized Combining Forms in ESP: Stable and Unstable Formations

Specialized CFs are certainly expanding the lexicon of ESP with new words. The aim of this section is to establish the productivity of these morphological elements used in different areas and with diverse degrees of stability. Some of them are more stable and produce large sets of new words which deserve inclusion in lexicographic works, others, being more specific and technical, have a limited applicability, but still contribute to the lexical innovation of ESP.

4.1 Neoclassical Combining Forms

Neoclassical CFs mainly pertain to the areas of chemistry, biochemistry, and medicine, with rare exceptions.

Among initial CFs, *ichno-*, attested since 1956 and obtained from Greek ἴχνος 'track, trace', is used in palӨontology with reference to trace fossils, especially in terms relating to their taxonomy. In those terms, *ichno-* combines with either final CFs (e.g., *ichnography*, *ichnology*) or full words (e.g., *ichnofauna*, *ichnofossil*).

A few of the new words are hapaxes (e.g., *ichnofamily* and *ichnographic*(*ally*)), yet most of them are stable neologisms (e.g., *This confirms the capacity of the body fossil record and* **ichnology** *to complement each other.* NOW, 17/03/2023).

Another specialized CF coming from Greek *λεξικός* 'lexicon' is initial *lexico-*, which has been attested from 1953 in modern linguistic terms denoting 'lexical and −'. It often combines with adjectives (e.g., *lexico-behavioural*, *lexico-grammatical*), but especially occurs before nouns (e.g., *lexico-dynamics*) or final CFs (e.g., *lexicography*). Nouns such as *lexicogrammar*, *lexicology*, *lexico-statistics* occur more frequently than the derivatives *lexico-grammatical*, *lexicologist*, *lexicostatistical*, *lexico-statistician*, with only one occurrence in either COCA or NOW. Nevertheless, many of these words (e.g., *lexicographer*, *lexicography*) are amply attested in dictionaries and corpora (e.g., *Davis studied at Rhodes University and Oxford before working in* **lexicography** *at the Oxford English Dictionary.* NOW, 07/01/12).

In chemistry, *seco-* (recorded since 1951) is a formative element used in naming derivatives, especially of steroids, in which fission of a ring has occurred, as in *seco-dicarboxylic* and *seco-steroid*. The origin of the CF is from Latin *secāre* 'to cut', followed by an *-o-* connective (Note 3). The use of this initial CF is very scanty, though, as the only word attested in corpora is *seco-steroid* (e.g., *The body metabolizes it into a* **seco-steroid** *hormone which, like all steroids, regulates genetic transcription, a basic and profound mechanism of action unlike any other vitamin*. NOW, 14/06/18).

The CF *synapto-*, from Greek *συναπτ-ικός* 'connective', has been recorded in the OED from 1962. It is used as CF of 'synapse' in various terms in physiology: e.g., *synaptology*, *synaptosomal*, and *synaptosome*. Frequent formatives in corpora include *synaptogenesis*, *synaptonemal*, *synaptophysin*, *synaptotagmin*, whereas *synaptoclastic*, *synaptogenic*, *synaptologist*, *synaptoplasticity*, and *synaptotoxicity* occur only once either in COCA or in NOW (e.g., *Ongoing research supports psilocybin's ability to stimulate* **synaptogenesis** *and neuroplasticity in the brain*. NOW, 09/05/23).

Final CFs are more frequent in corpora than initial CFs in terms of both types and tokens. In chemistry and biochemistry, the final CF *-mer* (from ancient Gr. *-μερής* 'part, segment', first attested in 1975 in the OED) forms nouns denoting particular kinds of polymer (as in *dimer*, *oligomer*) or isomer (as in *epimer*, *tautomer*). It can also be suffixed to numerals (in full form or more usually as digits) to form nouns denoting polymers consisting of a given number of units (as *9-mer*, *16-mer*). Combinations with *-mer* mainly include stable words such as *copolymer*, *monomer*, *polymer*, *tetramer*, *trimer* (e.g., *This implied that the complex consisted of three identical subunits, arranged as a symmetrical* **trimer**. COCA, 01/08/18), with rare nonce words or hapaxes such as *25-mer* and *36-mer*.

Another frequent final CF is *-ogen*, first attested in 1961 in the OED and coming from an *-o-* connective and the Greek root *γίγνεσθαι* 'be born, become'. In biochemistry, it forms the names of an inactive precursor of a compound to whose name it is appended, as in *hypertensinogen*, *kininogen*, etc. English has inherited from French the names *hydrogène* and *oxygène*, which were adopted into English with the ending *-gene*, afterwards altered to *-gen*. On the analogy of these words, a considerable number of new terms have been added to the English vocabulary of chemistry, in which the ending *-(o)gen* expresses the sense 'that which produces'. They are usually names of chemical substances, as in *amidogen*, *cyanogen*, *nitrogen*, etc., or of classes of substances, as in *halogen*. The most frequent words in both corpora are *hydrogen*, *estrogen* (e.g., *The basic components of HRT are* **estrogen** *and progestin, which is a form of progesterone*. NOW, 10/11/23) and *nitrogen*, but novel nonce words such as *angiogen*, *astrogen*, *carbogen*, *entactogen*, *vitrogen* also demonstrate the profitability of *-ogen* in the coinage of new vocabulary.

The CF *-penia*, first attested in 1971 in the OED, originates from ancient Greek *πενία* 'poverty' and is primarily used in medicine. From the end of the nineteenth century, it has been found in a small number of English formations and adaptations of foreign words, apparently first in *leucopenia*, and later in *granulocytopenia* and *neutropenia*. It chiefly combines with first elements of Greek origin, as in *leukopenia*/*leucopenia*, *osteopenia*, *sarcopenia* (e.g., *Recent estimates suggest that* **sarcopenia** *affects 10% to 16% of the elderly population worldwide*. NOW, 13/10/23), and *thrombocytopenia*, which are stable formations, unlike the hapaxes *cytoglucopenia*, *erythroblastopenia*, *plancytopenia*, *reticulocytopenia*, and *vasopenia* attested only once in the corpora.

The CF *-valent*, from Latin *valentem* (present participle of *valēre* 'to be worth'), has been attested in the OED from 1977. It is a formative element occurring in a few words of general currency and in various scientific contexts, used with prefixes denoting number. For instance, in chemistry and immunology, it forms adjectives denoting 'having a valency of the specified number' (e.g., *trivalent*, *quadrivalent*, *tetravalent*). It is also used

with Arabic numbers prefixed (e.g., *14-valent*). While *ambivalent* (with the prefix *ambi-*), *covalent*, *multivalent*, and *quadrivalent* are amply attested in the corpora (e.g., *A **quadrivalent** vaccine works by stimulating an immune response against four different antigens, such as four different viruses or other microorganisms*. NOW, 09/02/23), attestations of *hexacovalent*, *hypovalent*, *omnivalent*, and *plurivalent* are rare occurrences.

Table 1 summarizes the quantitative data resulting from the COCA and NOW corpora showing the productivity of specialized neoclassical CFs in terms of type frequency, token frequency, and hapax legomena.

Table 1. Frequencies of occurrence of neoclassical CFs in the COCA and NOW corpora

	TYPE FREQUENCY		TOKEN FREQUENCY		HAPAX LEGOMENA	
	COCA	NOW	COCA	NOW	COCA	NOW
ichno-	3	14	15	157	1	4
lexico-	14	12	188	1,803	3	1
seco-	1	1	3	1	0	1
synapto-	13	19	76	361	3	5
-mer	10	14	2,366	26,601	1	0
-ogen	38	83	24,198	145,067	5	16
-penia	14	29	380	3,135	2	5
-valent	20	43	2,990	13,529	3	7

The high frequency of CFs such as initial *lexico-* and *synapto-* and final *-ogen* and *-valent* shows the "realised productivity" or "extent of use" (Baayen, 1993) of such morphological elements and their profitability (Corbin, 1987; Bauer, 2001, p. 49) for the expansion of ESP lexicon. The fact that these are also used to coin nonce terms attested only once in corpora (hapaxes) signals their "expanding productivity", measuring the rate at which these morphological categories are expanding and attracting new members. Some neoclassical CFs (e.g., *ichno-* and *seco-*) are quantitively less productive, but their lower frequency is correlated to a sector-related use and specificity.

4.2 Abbreviated Combining Forms

Abbreviated CFs do not have a classical origin, but are obtained by shortening existing words and combining them with bases. Many of them are used to coin new words in the areas of digital technology, electronics, and economics.

Among initial CFs, one of the most frequent is *cyber-*. Shortened from *cybernetic*, it was originally (from 1961) used to create words, mainly nouns relating to computers, information technology, and virtual reality, or denoting futuristic concepts. Later, *cyber-* was specifically used to form terms relating to the Internet, probably influenced by analogy to *cyberspace*. Formations include established terms, such as *cyberart*, *cyberattack*, *cybercommunity*, *cybercrime*, *cyberfriend*, *cyberlover*, *cyber-romance*, *cybersecurity*, *cyberspeak*, *cybersphere*, *cyberterrorism*, and *cyberworld*, as well as temporary nonce words (e.g., *cyber-family*, *cyber-debate*, *cyber-piracy*, *cyber-protest*, *cyber-secure*) whose number is daily increasing. The following example illustrates *cybercrime* in context: *In this age of rapid technological advancement and widespread internet usage, **cybercrime** has emerged as a grave menace to our society, economy and personal security* (NOW, 23/11/23).

Shortened from the adjective *digital*, *digi-[1]* is an initial CF attested from 1960 to form nouns denoting 'a digital device, system, etc.', as in *digibox*, *digicam*. Later its meaning has been reinterpreted and, as a result, the initial CF has become secreted (cf. § 4.3). Besides *digibox* and *digicam* (e.g., *Film high DR roll-off in the highlights is sorely missed with **digicams**.* NOW, 21/10/19), corpora attest the use of neologisms such as *digicode*, *digicorder*, *digipack*, *digiPad*, *digitech*, as well as nonce words occurring only once (e.g., *digicell*, *digi-drum*, *digiPod*, *digi-spoon*, and *digitape*).

As announced, the initial CFs *e-[1]* and *e-[2]* are kept distinct because of their different origin and use. The initial CF *e-[1]* is earliest attested in 1969 with the meaning *electric*, from which it is abbreviated. With this sense, it forms nouns denoting vehicles powered by electricity instead of or as well as more traditional means, as in *e-car*, *e-bicycle* (or *e-bike*), and *e-scooter*. In corpora, only stable nouns occur, such as *e-boat*, *e-bus*, *e-cycle*, *e-ferry*, *e-jeep*, *e-power*, *e-rail*, *e-street*, *e-train*, and *e-van*, although *e-car* and *e-scooter* recur more frequently (e.g., *The launching of the **e-scooter** follows the commencement of the construction of Electrum's electric scooter factory in June 2023.* NOW, 23/11/23).

As a separate entry, the OED offers another initial CF *e-[2]*, more recent [1988] and this time abbreviated from

electronic. Found in numerous recognized formations from the late 1980s, as in *e-book* and *e-fit*, early formations are often preceded by corresponding compounds in *electronic* whose spelling is as two separate words (e.g., *electronic book* [1978] vs. *e-book* [1988], *electronic cash* [1967] vs. *e-cash* [1994]).

The CF *e-²* is indeed prefixed to nouns to indicate involvement in electronic media and telecommunications (esp. concerning the Internet), usually to distinguish objects or actions from their non-electronic counterparts. In the late 1990s and early 2000s, formations starting in *e-²* became very numerous and conspicuous in their frequency of use. The three specific semantic areas that these formations cover are: (1) terms relating to the publication or exchange of information in an electronic format (e.g., *e-edition*, *e-journal*, *e-publication*, *e-reader*, *e-text*, *e-zine*); (2) terms relating to computers and the Internet, i.e. computer-mediated and Internet-based (e.g., *e-fit* and other nonce words); and (3) terms relating to electronic financial transactions (e.g., *e-bill*, *e-cash*, *e-commerce*, *e-currency*, *e-dollar*, *e-money*, *e-ticket*). The latter more specialized meaning has been recorded from 1992. Besides *e-mail* (both noun and verb), the most common combinations are *e-book*, *e-cigarette*, *e-commerce* (e.g., *The business continues to deliver acceleration in **e-commerce** growth in the country.* NOW, 29/11/23), *e-learning*, *e-reader*, whereas nonce words in COCA include *e-product*, *e-progress*, *e-proof*, *e-quality*, *e-recruit*, *e-recycling*, *e-revenue*, *e-revival*, *e-reward*, *e-servicing*, *e-shopkeeper*, *e-smoker*, *e-smoking*, and *e-solution* among others.

The initial CF *econo-*, attested from 1964 in the OED, is found in a small number of formations. It is found earliest in *econometry/econometrics* 'the branch of statistical theory concerned with the analysis of economic phenomena' and related words, and in *econobox*, referring to 'an economy-sized box of something' or 'a nondescript economical car'.

As a CF, it can be shortened either from *economic* and prefixed to adjectives with the sense 'economic and –' (e.g., *econo-political*), or from *economy* and added to nominal bases. In the latter case, it can form nouns denoting 'things which are economical (or promote economy), inexpensive, or cheap' (e.g., *econocar*), or nouns denoting 'things or persons of a type involved in or associated with economics' (e.g., *econo-politics*). *Econometrics* and the related adjective *econometric* are the most frequent words in both corpora (e.g., *The popularity of **econometrics** shows the feelings of the people and entrepreneurs.* NOW, 16/09/23), yet several nonce words are also attested (e.g., *econocentric*, *econocratic*, *econography*, *economode*, *econopack*, *econoperformance*, *econo-size*, etc.).

The initial CF *lamino-* comes from the abbreviation of the adjective *laminal* 'produced by the blade of the tongue', with an additional *-o-*, and is used in the medical field. The first attested examples include *lamino-palatal* [1966] and *lamino-dental* [1968]. Other combinations found in corpora include *laminoplasty* (e.g., *Modified expansive open-door **laminoplasty** technique improved postoperative neck pain and cervical range of motion.* COCA, 01/03/19), *laminography*, *laminopathy*, and the nonce term *laminotomy*.

Abbreviated from *negative*, *nega-* was first attested in 1973. It is used to form the names of 'negative counterparts of things (chiefly units of measurement)' (e.g., *negawatt*), now especially in the context of energy (or other resources) saved as a result of conservation measures (e.g., *Basically, it said that RTOs, when organizing competitive auctions in wholesale markets, have to treat a megawatt of avoided energy use – a **negawatt** – the same way they treat a megawatt of energy.* NOW, 26/01/16). Examples of formatives attested in the COCA and NOW corpora include *negabarrel*, *nega-bobbing*, *negabot*, *negacaptain*, *nega-verse*, and the nonce words *nega-cancer*, *nega-carbon*, *negacoal*, *negaholism*, *nega-input*, *neganetwork*, *nega-oil*, *negatree*, and *Nega-Trump*.

The initial CF *petro-¹*, shortened from *petroleum*, first occurred in the mid-twentieth century in *petrochemical* [1942] and *petrochemistry* [1942], although *petromax* (from G. *Petromax*) [1929] 'a proprietary name for a type of vaporised paraffin pressure lamp' and *Petro-Forge* [1969] 'a kind of forging machine powered by a petrol engine' also belong here. As an abbreviated form, it obtains terms relating to petroleum, especially with reference to the political and economic power of oil-producing countries, as in *petropolitics* (e.g., ***Petropolitics** come with financial risks as well as upsides.* NOW, 04/01/23), *petro-power*, *petro-resources*, and *petro-wealth*. Formations abound in corpora, including, among others, both *petrochemical*, *petrography*, *petrologist*, *petrology*, *petromania*, *petropolitics*, *petro-state*, and a conspicuous number of nonce words (e.g., *petro-domination*, *petro-fueled*, *petro-military*, *petro-nation*, *petro-oil*, *petro-plastic*, *petro-supplier*, etc.).

The CF *porta-*, shortened from the adjective *portable*, apparently appeared earliest in *portapak* [1951] 'a portable system comprising a video camera and recording equipment'. It is used to form the (often proprietary) names of various movable or portable versions of manufactured objects, devices, structures, etc., as specified by the second element: e.g., *portacam*, *portacrib*, *Porta-John* (from slang *john* 'a toilet, lavatory'), *porta-kit*,

porta-office, *Porta-Phone*, *porta-printer*, *porta-screen*, etc. In corpora, we also find *portaball*, *portacabin* (e.g., **Portacabin** *schools were started by state government in 2012 to cater to tribals*. NOW, 19/01/23), *porta-cell*, *porta-chef*, *portal-gun*, *porta-potty*, and hapaxes such as *portabook*, *porta-box*, *portacomstyle*, *porta-filter*, *portaphone*, *portapool*, and *portaportal*.

The CF *syn-*, abbreviating *synthetic*, has been used from 1971 to form words denoting synthetic products, as in *syncrude* 'a synthetic product made from coal in imitation of crude oil', *synfuel*, *syngas*, *synjet*, *synoil*, and *synroc*. As such, it differs from the prefix *syn-*, a Latinized form of Greek *συν-* 'together' (e.g., *synergy*, *synthesis*), and generally carries stress. In corpora we mainly find stable formations, such as *syncrude*, *syndiesel*, *syn-free*, *syngas* (e.g., *Electrochemical splitting of CO_2 can create useful chemicals, such as* **syngas**, *Formic acid, Oxalic acid, and more*. NOW, 30/10/23), *synlawn*, *syn-turf*, and some rare nonce formations (e.g., *synsilk*, *synskin*).

The CF *-bot[1]* (from *robot*) is found in a number of formations from about 1966 onwards. The new words were originally blends involving *robot* (e.g., *mobot* from *mobile* and *robot*), but later *-bot[1]* was used to form nouns denoting 'a type of robot or automated device' or, in extended (metaphorical) use, to refer to 'a person regarded as an automaton', as in *fembot* [1976] 'a robot resembling a woman in appearance' and *nanobot* [1989] 'a nanorobot' (e.g., *The helical* **nanobot** *is made of silicon dioxide coated with iron*. NOW, 16/05/22) (cf. its specialized meaning in information technology, § 4.3). Other stable formations attested in corpora include *aerobot*, *autobot*, *chatbot*, *guardbot*, *medbot*, *microbot*, *minibot*, *newsbot*, *repairbot*, *sexbot*, *smartbot*, *teacherbot*, *twitterbot*, *waterbot*, whereas *agrobot*, *brainbot*, *girlbot*, *megabot*, and *securitybot* are some instances of nonce formations.

Attested in the OED from 1965, the abbreviated CF *-lect* (from *dialect*) is a terminal element used to designate a regional or social variety within a language, as in *idiolect* [1948] 'the linguistic system of one person, differing in some details from that of all other speakers of the same dialect or language'. It is also used in forming a number of technical terms in linguistics, as in *acrolect*, *basilect*, *isolect*, *sociolect*, etc. Hence, it has been converted to a noun and used to refer to 'a social variety of a language or dialect'. The most common formation in both corpora is *idiolect* (e.g., *Attackers can mimic the distinct* **idiolect** *of the target*. NOW, 18/11/23), followed by *Afrolect*, *basilect*, *ethnolect*, *interlect*, *multilect*, *sociolect*, while *hyperlect* and *topolect* are hapaxes.

The final CF *-olol* (from *propranolol*) is found in formations from the late 1960s onwards. It is typical of pharmacology, where it is used to form the names of 'beta-adrenoceptor blocking drugs derived from propranolol or having a similar molecular structure' (e.g., *acebutolol* 'a beta blocker which is partially selective for cardiac beta receptors, used in the treatment of hypertension, angina, and arrhythmia'). Specialized formations attested in corpora include *atenolol*, *bisabolol*, *bisoprolol*, *metabolol*, *metoprolol*, *nadolol* (e.g., *We aimed to test the effectiveness of isosorbide mononitrate as an adjunct to the b-blocker* **nadolol** *in the prophylaxis of first variceal bleeding in these patients*. COCA, 21/12/1996), *nebivolol*, *stanozolol*, *timolol*, while *nandolol*, *penbutolol*, *pindolol*, *practolol*, and *talinolol* are once-only occurrences.

In particle physics, *-onium* represents the abbreviation of *positronium*. Since 1987, it has been used to form the names of bound states of a particle and its antiparticle, as in *charmonium*, *nucleonium*, *toponium*. This is not to be overlapped with the homonymous antecedent form, occurring in chemistry from 1858 in *carbonium*, *hydrazonium*, *nitronium*, *phosphonium*, which instead comes from *ammonium* and forms the names of 'complex cations that contain a more or less electronegative central atom'. The only stable word attested in corpora is *charmonium* (e.g., *And, given that the specific quark is called "charm," its existence opens vast possibilities for puns – and that's without even getting into the fact that the technical term for the full family of particles containing these quarks is "charmonium"*. NOW, 14/07/20), while *quarkonium* is a once-only formation attested in NOW.

The CF *-ylidene*, from *ethylidene*, earliest attested in 1971, is used in chemistry as an adaptation of a French chemical name. It is used to form the names of divalent radicals in which both valencies derive from the same atom, as in *benzylidene*. It replaces the *-idine* suffix when the name of the parent compound does not end in *-yl* (cf. *alkylidene*, *propylidene*, etc.). Corpora attest stable words such as *2-chlorobenzylidene*, *4-methylbenzylidene* (e.g., *An order came into force on Wednesday banning lotions containing oxybenzone, octinoxate,* **4-methylbenzylidene** *camphor or butylparaben from Thailand's marine national parks*. NOW, 04/08/21), *benzylidene*, *cyclopentylidene*, *nitrobenzylidene*, *o-chlorobenzylidene*, *retinylidene*, but *2-xhlorobenzylidene* is attested only once in NOW.

Table 2 summarizes the quantitative data resulting from the COCA and NOW corpora showing the productivity of specialized abbreviated CFs in terms of types, tokens, and hapaxes.

Table 2. Frequencies of occurrence of abbreviated CFs in the COCA and NOW corpora

	TYPE FREQUENCY		TOKEN FREQUENCY		HAPAX LEGOMENA	
	COCA	NOW	COCA	NOW	COCA	NOW
cyber-	423	282	8,821	8,687	129	64
digi-[1]	20	12	208	412	5	20
e-[1]	8	24	160	20,422	0	0
e-[2]	279	281	62,061	639,728	19	0
econo-	41	31	873	2,851	21	88
lamino-	2	6	31	70	1	0
nega-	15	11	32	123	8	5
petro-[1]	66	87	1,025	31,263	37	212
porta-	19	36	221	1,951	7	16
syn-	6	7	206	3,873	2	0
-bot[1]	35	43	285	10,055	9	166
-lect	4	10	33	96	1	1
-olol	13	20	145	628	3	3
-onium	1	2	6	9	0	1
-ylidene	2	9	4	36	0	2

As Table 2 shows, specialized abbreviated CFs are highly frequent in corpora, particularly initial *cyber-* and *e-²*, productively forming hundreds of types, but also initial *petro-¹* and *econo-*, and final *-bot¹*, used to form both neologisms and nonce words. The high number of tokens occurring in corpora and the numerous hapaxes obtained through these morphological elements confirm the stability of the CFs and their acceptability for novel formations and vocabulary expansion.

4.3 Secreted Combining Forms

Secreted CFs involve both abbreviation and reinterpretation. Formally, they are obtained from the back- or fore-clipping of a word. Semantically, they involve either a generalization or a specification process, which contributes to conferring a certain level of abstraction on the CF. Many of them are used to coin new words in the areas of digital technology and, less frequently, economics.

An initial CF which is not only abbreviated (see its concrete use for digital devices in § 4.2) but also secreted is *digi-²*. Attested from 1986 as a secreted CF, it is used to create 'nouns denoting a product, process, person, etc., relating to or characterized by the use of digital technology', such as *digi-age*, *digi-art*, and *digi-novel*. On some occasions, it is also prefixed to the names of styles of popular music (as in *digi-funk*, *digirock*, etc.) to form 'nouns denoting music incorporating sounds generated or modified digitally'. Hence, unlike *digi-¹*, *digi-²* does not refer to devices, but to more abstract entities such as art, age, or music, as in *digisphere* (e.g., *You can just launch and then let it disappear into the **digisphere**.* NOW, 02/06/19). Other stable formations attested in corpora include *digibook*, *digiconomics*, *digiconomy*, *digimarketing*, *digi-paint*, *digiphile*, *digishop*, *digitour*, *digizine*, while *digidance*, *digidata*, *digi-friendly*, *digimusician*, *digirama*, and *digitask* are nonce words.

The formation of the secreted CF *m-*, from *mobile*, is clearly analogical: its similarity with *e-²* (see § 4.2) is marked formally – both retain only the initial letter of a word, being rather borderline with initialisms – and semantically. Recorded earliest in *m-commerce* [1997], *m-* forms terms relating to 'commercial activity conducted through mobile electronic media and devices, especially mobile phones', as in *m-banking*, *m-payment*, *m-ticket*, etc.

Since 2001 it has also been used to create terms (occasionally temporary words and ad hoc formations) relating to 'social and cultural activity or phenomena conducted through or using mobile electronic media and devices', as in *m-government*, *m-health*, and *m-voting*, but earliest in *m-learning*. The analogical nature of these formations is confirmed by the existence of *e-²* parallel formations (cf. *e-commerce*, *e-learning*) (e.g., *Share of **m-commerce** into e-commerce is projected to reach 80 percent by 2024.* NOW, 29/11/23). Other recognized formations are attested in the COCA and NOW corpora: e.g., *m-cash*, *m-enterprise*, *m-finance*, *m-government*, *m-retailing*, *m-service*, *m-system*, *m-ticket*, while hapaxes in COCA include *m-banking*, *m-finance*, *m-healthcare*, *m-mailbox*, and *m-money*.

Besides being an abbreviated initial CF shortened from *petroleum* (e.g., *petrochemistry*, see § 4.2), *petro-²* is also secreted when it forms terms designating 'revenue, esp. foreign exchange, that derives from petroleum exports', as in *petrodollar* [1973] 'a notional monetary unit earned by a country from the export of petroleum', *petrocurrency* [1974] 'the currency of a petroleum-exporting country', *petrobillion*, *petro-naira*, and *petropound*.

The secretion process activates semantic specification in this case. The most common word in corpora is *petrodollar* (e.g., *Since the menace of armed robbery went full throttle in the immediate* **petrodollar** *Nigeria of the early 1970s, it has grown further into becoming a social pandemic today.* NOW, 07/05/23), yet other formations are attested: e.g., *petrobank, petro-billionaire, petro-currency, petroeuro, petromoney, petroyuan,* and the nonce terms *petro-bourgeoisie, petro-charged, petro-shilling, petro-wealthy,* and *petro-welfare*.

The initial CF *robo-*, clipped from *robot*, has been attested from 1988 in a number of formations with the sense 'a/an – resembling a robot, especially in being resilient, emotionless, or futuristic'. Examples include both recognized nouns, such as *robocop* 'a robotic or bionic law enforcement officer' (used in science fiction from 1957) and *robocall* [1998] 'an automated telephone call which delivers a recorded message' (e.g., *But on Sunday Khatari cut a* **robocall** *to voters urging them to vote for Kagan.* NOW, 06/11/23), and nonce words such as *robo-boxer, robo-candidate,* and *robodisco*. In corpora, we also find *robo-adviser/robo-advisor, robocaller, robocar, robocat, robodoc, robodog, roboform, robohelp, robonaut, robo-signer, robosigning,* etc. Nonce words are surprisingly frequent, a few instances being *robobike, robobox, robobug, robobuilder, robodoctor, robohunter, robokiller, robomom,* and *robostaff*.

Shortened from the same base *robot*, the final CF *-bot²* is instead used in computing and information technology to form nouns denoting 'a type of automated program or (Internet) software, especially one which searches out information' (cf. also abbreviated *-bot¹* in § 4.2). More established combinations of this nature are *infobot* [1986] 'any of various automated systems for providing or obtaining information', *knowbot* [1988] 'a program designed to search through large numbers of databases and retrieve information', and *cancelbot* [1993] 'a program that searches for and deletes specified postings from Internet newsgroups'. Corpora also attest *Applebot, biobot, chatterbot, eventbot, filebot, Googlebot, junkbot, shopbot, spambot, talkbot, Twitterbot,* etc.

Some of the less common or more ephemeral forms are *searchbot, spybot* (e.g., *Finish them, and grab the* **Spybot** *before sliding down to the bazaar.* NOW, 10/06/21), and *warbot*. A semantic specification process has intervened in the formation of this CF.

The final CF *-nomics* is found in formations from the first half of the nineteenth century in the sense denoting the science or study of a subject specified by the first element, as *pyronomics, phoronomics,* etc. However, from the second half of the twentieth century, it has formed nouns denoting (often semi-humorous) fields of *economics*, from which it is shortened, as specified by the first element. It is therefore reinterpreted, via a specialization process, as 'the economic policies of –', the second element often being the name of a president of the United States, as in *Nixonomics* [1969] (from Richard *Nixon*), *Reaganomics* [1970] (from Ronald *Reagan*), and *Clintonomics* [1992] (from Bill *Clinton*), or another famous leader or political authority, as in *Rogernomics* [1985] (from the name of *Roger* Owen Douglas, New Zealand Minister of Finance).

Other attested formations include *Abenomics, Berlusconomics, Bushonomics, Carternomics, Gandhinomics, Obamanomics, Osbornomics, Popenomics, Putinomics, Thatchernomics, Trumponomics* (Note 4), and the nonce terms *corruptonomics, Dixonomics, dole-nomics, Dollarnomics, Gorbanomics, Wikinomics, womanomics,* some of which extend the base to a common noun (i.e. *corruption, dollar, woman*). A very recent formation occurring 2,186 times in NOW is *Bidenomics* (e.g., *U.S. President Joe Biden touts the administration's "***Bidenomics***" agenda during an October speech... from the White House.* NOW, 22/11/23).

The final CF *-verse* (from *universe*) has been used from 1981 to form nouns denoting 'the sphere or realm of what is specified or indicated by the first element'. Before becoming a recognized CF, the blends *nulliverse* [1847] 'a world devoid of any unifying principle or plan' and *multiverse* [1895] 'the universe considered as lacking order or a single ruling and guiding power' occurred. Nowadays, *-verse* is frequently used in the context of computing and electronic communications, as in *blogiverse, digiverse,* and *Twitterverse* (e.g., *The* **Twitterverse** *further said that the airline just followed the rules.* NOW, 12/08/23). Other formations are attested in corpora: *Appleverse, megaverse, metaverse, negaverse, pluriverse, televerse, webverse,* etc. However, from 1993, it has also obtained nouns denoting 'the fictional world associated with a specified character, television series, author, etc.', as in *Potterverse* (from J. K. Rowling's *Harry Potter* series), *Honorverse* (from David Weber's *Honor Harrington* series), *Buffyverse* (from the black comedy fantasy film and later television series *Buffy the Vampire Slayer*), *Whoniverse* (from the television series *Doctor Who*), and the nonce terms *Bronteverse, Garfieldiverse, Marvelverse*. Therefore, the CF acquires more specific meaning in the secretion process.

Table 3 summarizes the quantitative data resulting from the COCA and NOW corpora showing the productivity of specialized secreted CFs. The absence of hapax legomena in NOW is generally related to the fact that the automatic search is limited to 1,000 results per CF.

Table 3. Frequencies of occurrence of secreted CFs in the COCA and NOW corpora

	TYPE FREQUENCY		TOKEN FREQUENCY		HAPAX LEGOMENA	
	COCA	NOW	COCA	NOW	COCA	NOW
digi-[2]	12	21	26	519	6	0
m-	9	23	38	2,643	5	0
petro-[2]	7	21	235	2,331	3	2
robo-	81	128	1,147	19,567	39	19
-bot[2]	7	12	117	551	2	0
-nomics	10	43	339	4,155	3	9
-verse	16	54	828	12,856	6	3

Table 3 indicates that secreted CFs are less frequent than neoclassical and abbreviated CFs, in that they represent a rather recent phenomenon, still ongoing and sometimes developing from blends or abbreviated forms. In most cases, indeed, shortening precedes the semantic reinterpretation process, as with *digi-[2]*, *petro-[2]*, and *-bot[2]*, all corresponding to antecedent abbreviated forms (see § 4.2). The scarce number of hapaxes also goes in the same direction, confirming that secreted CFs are contributing to the development of ESP less conspicuously than neoclassical and abbreviated forms. Nevertheless, some less specific CFs, like initial *robo-*, are producing both types and hapaxes, showing an increasing productivity and going towards stability and recognition.

5. Conclusions

The present study has investigated the role of combining forms in the development of ESP lexicon. Although all the CFs examined here are recorded in the OED, their productivity is increasing and the stability of the combining-form combinations resulting from them has been confirmed by corpus linguistic analysis. Corpus-based results have demonstrated that, besides neoclassical CFs already attested in ESP since the nineteenth century, abbreviated and secreted CFs are also currently contributing to the lexical expansion of specialized English.

Some of the CFs analyzed have formed series of new words, especially stable and recognized neologisms, but also nonce words or hapax legomena which are attested only once in corpora, thus showing the potential profitability of these morphological elements in the coinage of still novel formations.

CFs are transitional phenomena often originated from a blending process, or from a shortening process followed by semantic reinterpretation. Secreted CFs, which involve semantic specialization or generalization, represent the last step of the morphological evolution from blend splinters to productive morphological elements that are very close to affixes. Hence, they are less frequently found in corpora, but their expanding productivity should be monitored in order to show how relevant they are to the vocabulary expansion of ESP.

The cases of *net-* (from *Internet*) and *-flation* (from *inflation*) belong here, in that they are splinters originated from blends (e.g., *netiquette ← Internet + etiquette*; *stagflation ← stagnant + inflation*), but they are becoming productive in the formation of novel words, such as *slumpflation* and *netizen* or *Netscape* (Note 5). Thus, the two splinters are potential CFs to be recorded in dictionaries not as part of blends, but as separate entries (i.e. initial CF *net-* 'related to Internet' and final CF *-flation* 'increasing inflation accompanying the state of the economy'), in spite of the scarce number of novel formations found in corpora.

Hapaxes in corpora, indeed, are often attested only once because of their sector-relatedness and specificity, but tend to be recognized very soon by the community of experts who are familiar with the lexicon of ESP and to be recurrently re-used in professional contexts as a confirmation of their lexicalization and stability.

References

Amiot, D., & Dugas, E. (2021). Combining forms and affixoids in morphology. In R. Lieber (Ed.), *The Oxford Encyclopedia of Morphology*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.569

Arndt-Lappe, S., Braun, A., Moulin, C., & Winter-Froemel, E. (Eds.) (2018). *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity, and Ludicity*. Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110501933

Baayen, H. R. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 181–208). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-3710-4_7

Bauer, L. (1983). *English Word-formation*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165846

Bauer, L. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486210

Bauer, L. (2017). *Compounds and Compounding*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108235679

Bauer, L., Beliaeva, N., & Tarasova, E. (2019). Recalibrating productivity: Factors involved. *Journal of Word Formation*, *3*(1), 44–81. https://doi.org/10.3726/zwjw.2019.01.02

Booij, G., & Hüning, M. (2014). Affixoids and constructional idioms. In R. Boogaart, T. Colleman & G. Rutten (Eds.), *Extending the Scope of Construction-based Grammar* (pp. 77–106). Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110366273.77

Brinton, L. J., & Closs, T. E. (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615962

Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique* (2 Vols). Tübingen: Max Niemeyer Verlag. https://doi.org/10.1515/9783111358383

Correia Saavedra, D. (2013). *Lexical Blending in English: From Splinters to Combining Forms*. Unpublished Master Thesis. Faculty of Humanities, University of Lausanne.

Correia Saavedra, D. (2016). Automatically identifying blend splinters that are morpheme candidates. *Digital Scholarship in the Humanities*, *31*(1), 55–71. https://doi.org/10.1093/llc/fqu059

Kastovsky, D. (2009). *Astronaut*, *astrology*, *astrophysics*: About combining forms, classical compounds and affixoids. In R. W. McConchie, A. Honkapohja & J. Tyrkkö (Eds.), *Selected Proceedings of the 2008 Symposium on New Approaches in English Historical Lexis (HEL-LEX 2)* (pp. 1–13). Somerville, MA: Cascadilla Proceedings Project.

Mattiello, E. (2022). *Transitional Morphology. Combining Forms in Modern English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009168274

Oxford English Dictionary Online. (1989-). 2nd/3rd Edn. Oxford: Oxford University Press. Retrieved from http://www.oed.com/

Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841323

Prćić, T. (2005). Prefixes vs initial combining forms in English: A lexicographic perspective. *International Journal of Lexicography*, *18*(3), 313–334. https://doi.org/10.1093/ijl/eci026

Prćić, T. (2008). Suffixes vs final combining forms in English: A lexicographic perspective. *International Journal of Lexicography*, *21*(1), 1–22. https://doi.org/10.1093/ijl/ecm038

Stevens, C. M. (2005). Revisiting the affixoid debate. On the grammaticalization of the word. In T. Leuschner, T. Mortelmans & S. de Groodt (Eds.), *Grammatikalisierung im Deutschen* (pp. 71–83). Berlin/New York: De Gruyter. https://doi.org/10.1515/9783110925364.71

Warren, B. (1990). The importance of combining forms. In W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer & J. R. Rennison (Eds.), *Contemporary Morphology* (pp. 111–132). Belin/New York: De Gruyter. https://doi.org/10.1515/9783110874082.111

Notes

Note 1. The *Corpus of Contemporary American English* contains more than one billion words (more than 25 million words each year 1990–2019) from eight genres including spoken language, fiction, popular magazines, newspapers, academic texts, TV and film subtitles, blogs, and other web pages. Retrieved from https://www.english-corpora.org/coca/

Note 2. The *News on the Web Corpus* contains 18.3 billion words of data from web-based newspapers and magazines from 2010 to the present time. More importantly, the corpus is constantly updated and grows by about two billion words each year. Retrieved from https://www.english-corpora.org/now/

Note 3. See Plag (2003, pp. 155–159) for the medial *-o-* that often appears in "neoclassical compounds".

Note 4. Occurring with different spellings, namely, *Trump(-)onomics*, *Trumpenomics*, *Trumpnomics*, and *Trumpanomics*.

Note 5. The OED records the following blends: *netiquette* [1982] 'an informal code of practice regulating the behaviour of Internet users when using email, bulletin boards, chat rooms, newsgroups, etc.', *netizen* [1984] 'a person who uses the Internet, esp. habitually', *Netscape* [1988] 'a proprietary name for: a browser used to access and display documents on the World Wide Web and other computer networks'; *stagflation* [1965] 'a state of the economy in which stagnant demand is accompanied by severe inflation', *slumpflation* [1974] 'a state of economic depression in which decreasing output and employment in industry are accompanied by increasing inflation'.

Copyrights