Volume (26), No. (1) , e2026460 - (28/07/2026)

The Rise of Social Media and the Introduction of New English Lexemes: Insights into Productivity

الدور المتزايد لوسائل التواصل الاجتماعي بإدخال كلمات معجمية إنجليزية جديدة: رؤى حول الإنتاجية

Rose Al-Hawamdeh, Salaam Alhawamdeh, Areej Al-Hawamdeh, Haitham AlYousef, Mouad Al-Natour
Hussein Technical University; Jerash University
rose.hawamdeh@htu.du.jo, salaam.alhawamdeh@jpu.edu.jo, areejalhawamdeh9@gmail.com, Haythamhmo84@yahoo.com, 1b2otnon@gmail.com

*(Corresponding author) e-mail: rose.hawamdeh@htu.du.jo

الملخّص

تتناول الدراسة الحالية عدداً من الكلمات التي تم إدخالها حديثًا في اللغة الإنجليزية، تحديداً الإنجليزية البريطانية والأمريكية (من التسعينات إلى اليوم)، نتيجة للتأثير المتزايد للتكنولوجيا، وخاصة وسائل التواصل الاجتماعي. الكلمات التي يتم دراستها مأخوذة من ثلاث منصات وسائل تواصل اجتماعي مختلفة، وهي: 1) المدونات الإلكترونية، 2) تويتر، و3) فيسبوك. يتم اختيار الكلمات المحددة من خلال مراقبة المواقع الإلكترونية المعنية، والاستشارة مع القاموس الحضري على الإنترنت، والحوار غير الرسمي مع المتحدثين الأصليين للغة الإنجليزية الذين يتمتعون بمعرفة تكنولوجية. إذا جاز التعبير. ثم يتم تحليل هذه الكلمات من حيث أصلها وتطورها، وتواترها ومن ثم إنتاجيتها من خلال فحص قاعدة البيانات المقصودة على الإنترنت، مثل COCA وBNC وGloWbE، وسياق استخدامها ومدى توافقها مع وحدة الطبقات المعجمية للغة الإنجليزية القياسية. تبين أن العديد من هذه الكلمات الجديدة لم تظهر بعد في مثل هذه القاعدة من البيانات، على الرغم من انتشارها عبر الإنترنت. الكلمات المختصرة والمجزأة لها أكبر تواتر عام عبر قواعد البيانات الثلاث، تليها الكلمات المدمجة ثم المركبة. كما توجد كمية أكبر من الأشكال التركيبية مقارنة بالأشكال المعجمية. الكلمات المدمجة مع اللواحق هي القاعدة الأكبر أيضًا من حيث نوع تشكيل الكلمات، تليها المركبات. بشكل عام، على الرغم من أنها غير رسمية ومبتكرة للغاية، إلا أن المصطلحات الجديدة التي يتم مناقشتها تتوافق مع وحدة الطبقات المعجمية للغة الإنجليزية القياسية.

ABSTRACT

The current study examines a number of lexemes that have been newly introduced into the English language; specifically British and American English (from the 1990's to the present day), as a result of the increasing influence of technology, specifically social media. The words under investigation originate from three different social media platforms; namely: 1) weblogs; 2) Twitter; and, 3) Facebook. The selection of the particular words is done through observation of the online sites under examination, consultation with an online urban dictionary and informal dialogues with native speakers of English who are tech savvy, so to speak. These words are then analyzed for their etymology and development, frequency and hence productivity through examination of online corpora, namely COCA, BNC and GloWbE, context of use and conformity to the lexical strata uniformity of Standard English. It is found that many of these newly formed words are yet to appear in such corpora, although widespread across the internet. Clippings and splinters have the largest overall frequency across the three corpora, followed by blends and then, compounds. There is also a larger amount of compositional than lexicalized forms. Blends with affixes are further, the largest group in terms of word formation type, followed by compounds. In general, although very informal and innovative in nature, the novel terminology under discussion conforms to the lexical strata uniformity of Standard English.

Article history:

Submission ID: 460

Submission Date: 03/09/2025

Reviewing Date: 13/10/2025

Acceptance Date: 12/06/2026

Publishing Date: 28/07/2026

DOI: 10.6520/jrs.v26i1.460

Keywords:

Corpora, frequency, , Productivity, morphophonemic, social media, word formation

Cite as:

Al-Hawamdeh, R. et al. (2026) The Rise of Social Media and the Introduction of New English Lexemes: Insights into Productivity. Jersah for Research and Studies 26(1). https://doi.org/10.6520/jrs.v26i1.460e202660

© The authors (2025). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC), -which permits non-commercial re use, distribution, and reproduction in any medium, provided the original work is properly cited. for commercial re-use, please contact admin@jpu.edu.jo.

The Rise of Social Media and the Introduction of New English Lexemes
Insights into Productivity

Rose Fowler Al-Hawamdeh[1], Salaam M. Alhawamdeh[2], Areej Mohammad Mahmoud Al-Hawamdeh[2], Haitham M.K AlYousef[2], Mouad Mohammed Al-Natour[2]

[1]Al Hussein Technical University, Jordan
[2]Jarash University, Jordan

Corresponding author: Rose Fowler Al-Hawamdeh, rose.hawamdeh@htu.du.jo

Abstract

The aim of the current study is to examine a number of lexemes that have been newly introduced into the English language; specifically British and American English (from the 1990's to the present day), as a result of the increasing influence of technology, specifically social media. The 1990's was chosen as a starting point given this is when social media first started to emerge. The words under investigation originate from three different social media platforms: weblogs; Twitter, and Facebook. The selection of the particular words (32 in total) is conducted through observation of the online sites under examination, consultation with an online urban dictionary and informal dialogues with native speakers of English who are tech savvy, so to speak. These words are further analyzed for their etymology and development, frequency and hence productivity through examination of online corpora, namely COCA, BNC and GloWbE, context of use and conformity to the lexical strata uniformity of Standard English. The results reveal that many of these newly formed words are yet to appear in such corpora, although widespread across the internet. Clippings and splinters have the largest overall frequency across the three corpora (59.9%), followed by blends (30.7%) and then, compounds (9.2%). There is also a larger amount of compositional than lexicalized forms. Blends with affixes are further, the largest group in terms of word formation type, followed by compounds. In general, although very informal and innovative in nature, the novel terminology under discussion conforms to the lexical strata uniformity of Standard English.

Keywords: Corpora, Frequency, Productivity, Morphophonemic, Social Media, Word Formation.

الدور المتزايد لوسائل التواصل الاجتماعي بإدخال كلمات معجمية إنجليزية جديدة رؤى حول الإنتاجية

روز فاولر الحوامدة[1]، سلام محمد الحوامدة[2]، أريج محمد محمود الحوامدة[2]، هيثم اليوسف[2]، معاذ محمد النطور[2]

[1]جامعة الحسين التقنية، الأردن

[2]جامعة جرش، الأردن

المؤلف المراسل: روز فاولر الحوامدة، rose.hawamdeh@htu.du.jo

الملخص

تتناول الدراسة الحالية عددًا من الكلمات التي تم إدخالها حديثًا في اللغة الإنجليزية؛ تحديدًا الإنجليزية البريطانية والأمريكية (من التسعينات إلى اليوم)، نتيجة للتأثير المتزايد للتكنولوجيا، وخاصة وسائل التواصل الاجتماعي. الكلمات التي يتم دراستها مأخوذة من ثلاث منصات وسائل تواصل اجتماعي مختلفة؛ وهي: 1) المدونات الإلكترونية؛ 2) تويتر؛ و3) فيسبوك. يتم اختيار الكلمات المحددة من خلال مراقبة المواقع الإلكترونية المعنية، والاستشارة مع القاموس الحضري على الإنترنت، والحوار غير الرسمي مع المتحدثين الأصليين للغة الإنجليزية الذين يتمتعون بمعرفة تكنولوجية، إذا جاز التعبير. ثم يتم تحليل هذه الكلمات من حيث أصلها وتطورها، وتواترها ومن ثم إنتاجيتها من خلال فحص قاعدة البيانات المقصودة على الإنترنت، مثل COCA وBNC وGloWbE، وسياق استخدامها ومدى توافقها مع وحدة الطبقات المعجمية للغة الإنجليزية القياسية. تبين أن العديد من هذه الكلمات الجديدة لم تظهر بعد في مثل هذه القاعدة من البيانات، على الرغم من انتشارها عبر الإنترنت. الكلمات المختصرة والمجزأة لها أكبر تواتر عام عبر قواعد البيانات الثلاث، تليها الكلمات المدمجة ثم المركبة. كما توجد كمية أكبر من الأشكال التركيبية مقارنة بالأشكال المعجمية. الكلمات المدمجة مع اللواحق هي القاعدة الأكبر أيضًا من حيث نوع تشكيل الكلمات، تليها المركبات. بشكل عام، على الرغم من أنها غير رسمية ومبتكرة للغاية، إلا أن المصطلحات الجديدة التي يتم مناقشتها تتوافق مع وحدة الطبقات المعجمية للغة الإنجليزية القياسية.

الكلمات المفتاحية: قاعدة بيانات، التردد، الإنتاجية، علم الصرف الصوتي، وسائل التواصل الاجتماعي، تكوين الكلمات.

Introduction

The current study inspects various newly introduced (1990's onwards) British and American English lexemes given the expanding technological influence on language, especially with regards to social media. The rational behind choosing the 1990's as a starting point was because this was the period when social media first started to emerge. Weblogs, Twitter and Facebook were the social media platforms used to examine the corpora (32 words in total), as well as use of an online urban dictionary and communication with tech savvy native speakers of English. Analysis of the words under examination focused on their etymology and development, frequency and hence productivity through COCA, BNC and GloWbE. The context in which words were used and conformity to Standard English lexical strata was also examined. The Digital Revolution as defined by Schoenherr (2004) is:

> The shift from mechanical and analogue electronic technology to digital electronics which began anywhere from the late 1950s to the late 1970s with the adoption and proliferation of digital computers and digital record keeping that continues to the present day. (p.1)

Unreservedly, the expression also refers to the extensive modifications that have emerged as a result of digital computing and communication technology throughout and subsequent to the later stages of the 20th century. As Debjani (2014) argues "the Digital Revolution marked the beginning of the Information Age" (p.107), resulting in the mass production and pervasive use of "digital logic circuits, and its derived technologies, including the computer, digital cellular phone, and the internet" (ibid). Social media are computer-mediated technologies that assist the construction and distribution of information, thoughts, vocational awareness, and "other forms of expression via virtual communities and networks" (ibid). As a result, a phenomenal amount of new terminology has surfaced in the English language to incorporate this new online world, with evermore innovative ways of describing it. Social media platforms are certainly no exception. Caruntu-Caramen (2022) further adds:

> The dynamism of a language is sustained by the flexible and receptive character at the lexico-semantic level. An important contribution to the development and enrichment of the current vocabulary is made not only by the lexemes of English origin (Anglicisms), but also by the derivatives produced from them. (p.32)

In relation, the purpose of the current study is to examine a specific subset of these newly coined expressions and the extent to which they have influenced modern-day English.

Aims of the Study

The study aims to examine:

1) The origins of the technical words under discussion, as well as establishing what type of word formation they are such as blends, compounds, clippings, etc.
2) How the technical words under discussion have been developed through derivation; the process of forming a new word from an existing word, often by adding a prefix or suffix.
3) The frequency and hence productivity of the technical words under discussion, according to three corpora: namely; COCA, BNC and GloWbE.
4) Some of the contexts in which the technical words under discussion are used by native speakers of English.
5) Whether these new word forms violate the traditional lexical strata principles of Standard English.

Literature Review

Frequency

Language changes and evolves in its usage and one of the important sources of language change is frequency. According to Haspelmath and Simms (2010):

> The ways in which speakers use language have a profound influence on language structure, and frequency is one of the most important sources for system-external explanation of language structure (p.265).

Frequency affects the content of the lexicon, productivity and word class (ibid). It also signals language change, since it sheds light on the lucidness of language structure. Furthermore, frequency leads to leveling which in turn, leads to stabilization. On the other hand, it can also lead to disorganization. If for example, there is a disproportionate number of particular words, the morphological structure of a word may be affected which is a further signal of change. It is suggested that language systems move from chaos (i.e., flux) to stabilization (i.e., systematization). Hence, leveling eliminates morphophonological alternation, which is suggested to be problematic. Frequency in affect causes less phonological change, i.e., it acts as a decelerating function. For example, certain linguistic phenomena which are of a frequent nature, such as *go - went*, are immune to analogical leveling. It will always be the 'odd one out' as it were, i.e., the opposite of systematization. Moreover, on a phonological level, frequency results in speedier pronunciation of such words. It leads to phonological reduction because it is frequent; speakers take short cuts, e.g., the reduction from *do not* to *don't* (ibid).

Productivity

Emayakre et al. (2022) explains that "productivity as a morphological process is very pertinent because through it the native speakers are able to create new lexemes in their language" (p.16). This entails those grammatical processes are compared for their frequency in use, whether they be less frequently used or more so; this is to the extent that they lean towards lexicalization. Haspelmath (2024) goes further by stating:

> Productivity (the possibility of creating novel forms) is often mentioned as characteristic of inflectional patterns, and it is indeed a necessary prerequisite of completeness (p.56).

As Fernández-Domínguez (2010) suggests "productivity, first, bears a significant relationship with frequency because both seem to be subtly interlinked through low-frequency items". In general, productivity may be tested through identifying which grammatical forms are more likely to be exploited in the coining of new words (Haspel & Smith, 2010, p.114). In this case, word coinage is inclined to be converted to other forms solely by incorporating productive processes (ibid). In Standard English, for example, the formation of preterite and past participle forms of verbs by means of ablaut (cf. *sing–sang–sung*), is no longer deemed to be productive. Verbs in modern day English that are newly coined, to a large degree, use the 'weak' (regular) ending *-ed*. for the past tense and past participle seen (as is the focus of this study) in new technological verbs such as *spammed* and *e-mailed* (Harald & Lieber, 1991).

A few forms are hardly ever productive, whereas additional structures are perhaps employed by archetypal native speakers on a regular basis. In contrast, other processes (especially of a syntactic nature) might well be utilized productively on numerous everyday occasions. More than one correspondingly functioning pattern may also be comparably productive, to the extent that a speaker may face the dilemma of which form to make use of. For example, even a native

speaker may wonder whether it is potentially more appropriate to articulate colors or tastes pertaining to 'raisins' as *raisinish*, *raisiny*, *raisinlike*, or even *raisinly* (ibid).

Blending

'Affixation' is the process of adding a morpheme or affix to a word to create either a different form of that word or a new word with a different meaning which may also change the word category (Gries, 2014). With the rise of technology and concepts not previously in existence, different types of affixations have seen a phenomenal rise in the 21st century and beyond. 'Blending' (otherwise known as 'portmanteau', 'telescoping', 'lexical interlocking', and 'semantic conflation') for instance, "is the creation of new words by putting together parts of words that are not themselves morphemes" (ibid, p.642).

Such affixation includes bases already present in English, but to which affixes have been added to. In some cases, the meaning of the base does not fully (or completely) correspond to the newly formed words as a result of affixation. For example, from the noun *parent* (i.e. one that begets or brings forth offspring), we can form the gerund noun *parenting* (i.e. the raising of a child by its parents). With the phenomenal rise in the use of Facebook globally, and a transcultural desire to share every last mundane detail of one's life with others electronically, has arisen the lexeme *sharenting* (i.e., to share every detail of your parenting experience on Facebook). Now, whereas the word *share* means (have a portion of (something) with another or others), the term *sharenting* cannot be fully derived from the base and is in fact a blend of both *share* and *parenting*. Hence, such examples are non-compositional, i.e., are not predictable from the sum of their parts and are referred to as 'lexicalised'. In contrast, compositional compounds or blends consist of components which bear identifiable elements of their meaning as a whole (Cacciari & Padovani, 2012).

Blending is certainly a growing phenomenon within the realm of technical vocabulary, some newly formed lexemes of which are not entirely derived from already established bases, but from other newly formed lexemes (ibid). For instance, *vlog* (i.e., a blog in which the postings are primarily in video form) is derived from a fairly new (1990s) lexeme *blog* (i.e., a regularly updated website written in an informal style, abbreviated from *weblog*) and an established lexeme (1930s) *video* (i.e., a recording of moving visual images made digitally or on videotape, from Latin *videre* 'to see'). In this case, we also have an example of 'clipping' (a means of creating new words by shortening already existing words).

There are many types of blending depending on how they are formed. Algeo (1977) proposes dividing blends into three groups, namely: 1) phonemic overlap- a syllable or part of a syllable is shared between words, e.g., *chexting*, i.e., *cheating* and *texting*; 2) clipping - the shortening of two words and then compounding them, e.g., *weblog* to *blog*; and, 3) phonemic overlap and clipping - the shortening of two words to share a syllable and then compounding them, e.g. *newscast*, i.e. *news* and *broadcast.* Nevertheless, this taxonomy of blends is not typical amongst the categorizations provided by all linguists. Most blends are fashioned by the following methods: 1) the beginning of one word is added to the end of another (portmanteau'). For example, *simulcast* is a blend of *simultaneous* and *broadcast*; 2) the beginning of two words is combined. For example, *cyborg* is a blend of *cybernetic* and *organism*; 3) two words are blended around a common sequence of sounds ('overlaps'). For example, *motel* is a blend of *motor* and *hotel*; 4) multiple sounds from two component words are blended, while mostly preserving the sounds' order. For example, *slithy* is a blend of *slithe* and *slimy* and *stagflation* is a blend of *stagnation* and *inflation*; and, 5) a full word followed by or preceding a word part (otherwise known as a 'splinter'), as in *'motorcade'*, i.e., *motor* and *cavalcade* (ibid).

Acronyms / Initialisms

Other new word formations may appear in the form of acronyms / initialisms which refer to the first letters of words that make up a name or a phrase that are used to create a new word, such as WWW (World Wide Web). Both acronyms and initialisms are abbreviations, but there is a key difference between the two. Acronyms are abbreviations which are formed from letters of other words (usually the first letter of each word, although this is not always the case). The essence of the meaning of acronyms is that the consequential abbreviation should be pronounceable as a word, such as *RAM* (Random Access Memory); *LASER* (Light Amplification by Stimulated Emission of Radiation); *NASA* (National Aeronautics and Space Administration), and *OPEC* (Organization of Petroleum Exporting Countries) (Hiskey, 2012). Initialisms similarly, consist of letters of a particular name or phrase, which is typically the first letter of each word. However, unlike acronyms, the abbreviated forms of initialisms are not pronounced as words, but rather as individual letters. Examples include *FBI* (Federal Bureau of Investigation), *CIA* (Central Intelligence Agency), and *DVD* (Digital Video Disk). On the other hand, an abbreviation is simply a shortened form of a word, like *e.g.* (i.e. 'for example'), which is neither an acronym nor an initialism (ibid).

Coinage

Kosur (2012) purports that:

> Coinage is the word formation process in which a new word is created either deliberately or accidentally without using the other word formation processes and often from seemingly nothing (p.1).

'Coinage', also referred to as 'neologisms', (i.e. new word) is the process of making up entirely original words such as *Google* which was originally invented by Milton Sirotta, nephew of mathematician Edward Kasner in 1938, during a discussion of large numbers and exponential notation. *Google* is not only the name of a search engine but has now transformed to become a verb (i.e. search for information about (someone or something) on the internet using the search engine Google) or these days, any other search engine for that matter. The creation of novel words in English, not only relates to coinage, but also other "related word formation processes" (ibid) such as nonce words, borrowing (i.e. the process by which a word from one language is adapted for use in another), and calquing (i.e. a word or phrase borrowed from another language by literal, word-for-word or root-for-root translation). It is interesting that numerous coinages originate as ordinary brand names for commonplace items and products, such as *Kleenex* in the U.K., or *Fine* in some Arab countries, for facial tissues. Moreover, in some Arab countries specifically, coinages of English words for everyday items are common; even those which are not adopted in English speaking countries. Examples include *Nescafe* to refer to all instant coffee, *Pampers* to refer to all nappies, and *Cornflakes* to refer to all types of breakfast cereal.

Nonce Words

'Nonce' (i.e. for a single occasion) words are innovative expressions fashioned through the course of numerous and varied word formation processes with the consequential word is suitable for a particular lexical requirement which is not anticipated to reappear. For instance, the word *sandwich* derives from the 18th century *Earl of Sandwich* who first insisted on having bread and meat together while gambling, and *Hoover* originates from the *Hoover Suction Sweeper Company* which produced the first vacuum cleaner. Other sources include name of a person or a place. They are the 'eponyms' (i.e. a name or noun formed after a person) or 'antonomasia' (e.g. the substitution of an epithet or title for a proper name such as '*the Maid of Orleans'* for Joan of Arc, or the use of a proper name to express a general idea, such as *a Scrooge* for a miser (Kosur, 2012).

Lexical Strata

Phonology is the section of linguistics that concerns itself with:

> Sound regularities in languages: what sounds exist in a language, how those sounds combine with each other into syllables and words, and how the prosody (stress, accent, tone, and so on) of a language works (Lieber, 2009, p.168).

Phonology interrelates with morphology through numerous means. Morphemes may well have two or more diverse phonological forms whose manifestation could be entirely expected, or at least to a certain degree. For instance:

> Some phonological rules apply when two or more morphemes are joined together. In some languages morphemes display different phonological behavior depending on whether they are native to the language or borrowed into it from some other language (ibid).

In this study we also, to some extent, examine how phonology cooperates with morphology in the development of the technological words under discussion. Allomorphs are phonologically diverse alternates of an identical morpheme that "share the same meaning or function" (ibid). Distinctiveness refers to comprising of comparable, but not indistinguishable, sounds. For instance, the negative prefix *in-* in English is frequently pronounced as *in-* such as *intolerable*, yet in addition is on occasion pronounced as *im-* or *il-*, as in *impossible, illegal*. Since the entirety of these forms signify negativity, as well as attaching to adjectives in an identical fashion, they are considered as allomorphs of the negative prefix.

Regular past tense in English is a further example of allomorphy. While it is always spelled *-ed*, it has three variants in terms of pronunciation, i.e. [t] (*packed*), [d] (*bagged*), and [ɘd] (*waited*). Yet, given the difference in sounds, each one still allocates the past tense. As Lieber (2009) suggests:

> Building complex words is frequently accompanied by phonological effects such as assimilation or vowel harmony, yet in some languages such phonological effects do not apply uniformly across the entire lexicon of the language, but instead are confined to a subset of the lexicon. Indeed some languages have two or more different layers to their lexicons which behave differently in terms of phonological effects (p.168).

Corpora

Established in 1991, BNC (British National Corpus) is a 100-million-word text corpus of representative samples of written and spoken English composed from a wide range of resources. The corpus includes British English of the later 20th century from an extensive assortment of genres. The establishment of BNC originates from the alliance of three publishers; namely, Oxford University Press (as the main contributor), Longman and W. & R. Chambers, and the University of Oxford, Lancaster University, and the British Library. The corpus is limited to British English, with 90% written corpus samples, and the remaining 10% extracted from spoken language (British National Corpus).

COCA (Corpus of Contemporary American English) on the other hand, created by Mark Davies, Professor of Corpus Linguistics at Brigham Young University, is the largest corpus of American English that is openly accessible online. COCA includes an extensive assortment of texts from numerous genres (spoken, fiction, popular magazines, newspapers, and academic journals) and is composed of more than 450 million words from

more than 160,000 texts, 20 million words of which were accumulated between 1990 and 2015 (Corpus of Contemporary American English). The Corpus of Global Web-based English (GloWbE; pronounced as 'globe') enables its users to compare between different varieties of English, offering unequaled insight into different varieties of English, with approximately 1.9 billion words of text from twenty diverse countries (Corpus of Global Web-based English).

Sample, Methods of Data Collection and Analysis

The researchers collected a sample of 32 words relating to social media platforms that were introduced into the English language from the 1990's onwards. The words were divided equally into 3 categories in relation to the platforms they refer to, namely: 1) weblogs, 2) Facebook, and, 3) Twitter. The rationale behind choosing these categories is that they are representative sites of a vast amount of new English vocabulary. The selection of the particular words was done through observation of the online sites under examination, consultation with an online urban dictionary and informal dialogues with native speakers of English. These words were then analyzed for their etymology and development, frequency and productivity (through examination of online corpora), context of use and conformity to the lexical strata uniformity of Standard English.

Results and Discussion

Weblogs

Table (1) below presents 12 words which are derivations of the word *weblog* and their frequency according to COCA, BNC and GloWbE. A *weblog* is 'a regularly updated website or web page, typically run by an individual or small group, that is written in an informal or conversational style'. This term was invented in the 1990s with the onset of the internet and rise in online journalism, documented opinion and expression. If we divide this compound, we can analyze two nouns into: 1) *web* (n.), i.e.,' a complex system of interconnected elements' (Germanic), and, 2) *log* (n.), i.e. 'an official record of events' (Middle English).

Table (1). Frequencies of words relating to weblogs in COCA, BNC and GloWbE

No.	Word	Word Type	COCA	BNC	GloWbE		Total
					U.K.	U.S.	
1.	*blog*	(n. / v.)	5,968	0	7047	8707	21,722
2.	*blogosphere*	(n.)	278	0	986	1567	2831
3.	*microblogging*	(n. /v.)	50	0	244	143	437
4.	*vlog*	(n. /v.)	9	0	78	107	194
5.	*liveblogging*	(n. /v.)	2	0	76	31	109
6.	*moblogging*	(n. /v.)	1	0	3	2	6
7.	*blogonista*	(n.)	0	0	0	0	0
8.	*eco-blog*	(n. /v.)	0	0	0	0	0
9.	*mummy blog*	(n. /v.)	0	0	0	0	0
10.	*bloggable*	(adj.)	0	0	0	0	0
11.	*GRWM*	acronym	0	0	0	0	0
12.	*OOTD*	initialism	0	0	0	0	0

As Table (1) reveals, the most frequently used term in this group is *blog*, probably since this is a clipping of the original word *weblog*. This is followed by *blogosphere*, *microblogging*, *liveblogging*, *vlog* and *moblogging*. The sections below discuss some patterns found in the types of word formation relating to the words presented in Table (1) as well as their etymology, morphonemics, group systems relating to morphological phenomena, different strata and examples of context in which these terms are found. Examples of context were extracted from GloWbE, excluding those terms which had no frequency. For such terms, the authors provide other examples from the internet.

Compounds

The following words from the set presented in Table (1) above are compounds: 1) *microblogging*; 2) *liveblogging*; 3) *blogosphere*; and, 4) *mummy blog*, since they all contain morphologically complex words belong to two (or more) word families simultaneously. We discuss them each individually in more detail below. *Microblogging*, which can be either a verb or a noun, refers to 'writing very short blogs'. *Micro* (a Greek base) is an adjective which means 'small', *blog* is a clipping of weblog, and *ing* is a Germanic suffix 'denoting a verbal action / gerund'. The original term for these kinds of logs was *tumblelogs* (2005), but this was then adapted to *microblogging* (between 2006-2007) by mainstream internet providers. Consider the following example from 'You Don't Own Me: The Value of Virtual Property Beyond Bits and Bobs' (U.K.):

> *In the US, the state of Oklahoma in 2010 was the first to pass a bill decreeing that executors could gain access/control of the deceased's social networking,* ***microblogging*** *or email accounts -- a far cry from the Yahoo case in 2005 in which the parents of a dead marine were denied access to his emails.*
>
> (http://www.wired.co.uk)

Liveblogging, which emerged in the mid-2000s, and can also be a verb or a noun, refers to 'the action of commenting (or) commentary about (an event) while the event is taking place, typically in the form of frequent short blog updates'. *Live* is a Germanic adjective which means 'transmitted at the time of occurrence'. See the discussion for the clipping *blog* for the suffix *ing*. Consider the following example from 'US election briefing: The View from Nashville | World News' (U.K.):

> *But the new, negative tone of the race - however much McCain may be to blame for it - could pose problems for him in a one-on-one encounter with Obama, since it clashes with the personal image he has long sought to project. I'll be* ***liveblogging*** *the debate.*
>
> (http://www.guardian.co.uk/world)

Blogosphere which is a noun, was the most frequently apparent term after *blog* across the three corpora. It refers to 'the world of blogging' and resembles the older word *logosphere*, from the Greek *logos* and *sphere*, interpreted as *world*, i.e. 'the world of words'. This term was coined on September the 10th, 1999 by Brad L. Graham, as a joke but was soon absorbed into mainstream internet lingo. Consider the following example from the 'Public Option Is Not Symbolic; It Is Foundational | FDL Action' (U.S.):

> *So how come 15-20 Blue Dogs carry ten times the weight of 10-15 progressives? How come their thin wedge is so much more powerful than our thin wedge? Why does nobody outside of the progressive* ***blogosphere*** *seem to ask that question why?*
>
> (http://www. fdlaction.firedoglake.com)

Mummy blog (a noun), which refers to 'a blog covering parenting', although gaining momentum online from the 1990's onwards, was not found in any of the corpora utilized in this study. It consists of *mummy* (n.) (argued to be of Persian origin), the informal synonym for *mother* (Germanic) and *blog* (n.). An example of the use of this word is from 'How to Become a Mummy Blogger' (U.K.):

> *Want to make new mum friends, sound off about your parenting experiences, record your favourite moments, or seek help and advice. A* ***mummy blog*** *could be the answer.*

(http://www.madeformums.com)

Clippings and Splinters

Blog (n.) is a clipping of *weblog.* An example of the use of this word is from 'It Was A Good Day | Know Your Meme' (U.S.):

> *Further Research # On January 26th, 2012, the creator of the Murk Avenue Tumblr 6* ***blog*** *Donovan Strain published a post in which he claimed to have deduced the exact date described in the Ice Cube song as January 20th, 1992.*
>
> (http://www.knowyourmeme.com)

Vlog (n.) which emerged from the late 1990's, is a clipping and phonetic overlap of *video weblog* of the same meaning, made up of *video,* i.e., a recording of moving visual images made digitally or on videotape, from the Latin *vidēre*, and *log*, (again a clipping) from *weblog*, as discussed previously. An example of the use of this word is from 'She Could be One of Them | My Time as Mom' (U.S.):

> *Feeling a little less alone in this thing called mommyhood is a big one. Aramelle One Wheeler's World recently posted.* ***VLOG*** *The Big Cheese ##15662 Ann Krohn Rick: I'm happy to pay for your entitlements.*
>
> (http://www.mytimeasmom.com)

Moblogging (n. / v.) emerging from 2000 onwards, means 'blogging on the move, often using mobile phones to transfer the blogs to the internet'. *Mo* is a splinter of the Latin *mobile* (n.) i.e. 'mobile phone' (British English), *blog* (n.) is a clipping of *weblog* and the Germanic suffix *ing* denotes a verbal action, an instance of this, or its result. An example of the use of this word is from 'Serendipity in the age of online shopping - Flavrbox - Tummyrumble' (U.K.):

> *I first read a post titled "Whatever happened to serendipity? "At the time I was dabbling with "****moblogging****", an internet dinosaur word for "taking photos with my phone while out and about", something everyone today do all the time... #*
>
> (http://www.tummyrumble.hultberg.org)

Blends with Affixes

The following three words were also not found across the three corpora, but are used frequently online:

Bloginista is a noun which refers to 'a person who blogs about fashion.' It originates from *fashionista* (1990s), i.e., 'a devoted follower of fashion'. The word consists of *blog* and the informal English suffix *ista, '*forming nouns denoting a person associated with a particular activity', often with a derogatory force. Consider this example from the Facebook page of 'Visual Vamp':

> *I'm a bon vivant! Life is a visual banquet for all to share. Hug a* ***bloginista*** *today! This is a blog with original writing, and also a digest of the best of the posts on other decor and fashion blogs. If you see something that is yours, and you want credit, just tell me and I'll do it.*
>
> (https://www.facebook.com/Visual-Vamp)

Eco-blog (early 2000s), is a noun which depicts 'a blog covering environmental issues'. *Eco* is an English prefix relating to the environment, and the word *blog* is discussed above. Consider this example from 'Gypsey Soul' (U.K.):

> *Today I thought I'd list some of my favourite* ***Eco blogs*** *for you. In this list, you'll find blogs about zero waste, natural beauty tips, ethical living, seasonal recipes & lots more!*
>
> (https://www.gypsysoul.co.uk)

Bloggable (1990 onwards) is an adjective which means 'a subject which can be blogged about'. It consists of the noun *blog* and the suffix *able* meaning to 'be capable of', which normally occur in loanwords from Latin. Consider this example from "Bloggable and 'Scareware' added to Oxford Dictionary' (U.K.):

> *A slew of new words, including* ***bloggable*** *and scareware, have entered the Oxford Dictionaries for the first time because of the internet.*
>
> (https://www.telegraph.co.uk)

Acronyms

GRWM stands for 'Get Ready with Me' and emerged in the 1990's with the onset of weblogs. Vloggers show themselves getting ready online, whether it be for school, work or an outing for instance. It is an acronym because it is pronounced as one word - a pun for the growling sound some people make when they first get out of bed in the morning. Consider this example from 'Urban Dictionary' (U.K.):

> ***GRWM*** *First Day of Class! Morning routine, makeup, & outfit!"*
>
> (https://www.urbandictionary.com)

Initialisms

OOTD stands for 'Outfit of the Day' and also emerged in the 1990's with the onset of weblogs. It is usually used on fashion blogs, or other places where people show others what they are wearing or have worn. It is an intialism since each letter is pronounced separately, rather than as a single word. Consider this example from 'What Does OOTD Mean on Instagram? - He Spoke Style' (U.S.):

> *Photos of the* ***OOTD*** *variety are often selfies taken in a full-length mirror or a hastily snapped photo on the street with mobile phone's camera.*
>
> (https://www hespokestyle.com***)***

Twitter

Table (2) below presents 10 words which are derivations of the word *Twitter* and their frequency according to COCA, BNC and GloWbE. *Twitter* is an online news and social networking service where users post and interact with messages, called 'tweets'. These messages were originally restricted to 140 characters, but on November 7, 2017, the limit was doubled to 280 characters for all languages except Japanese, Korean and Chinese. We may suggest *Twitter* gets its name from the literal sense of the word, i.e., a series of short, high-pitched calls or sounds (late Middle English), but in relation to the actual site, i.e., in its new social networking sense, the term has a more metaphorical sense.

Table (2). Frequencies of words relating to Twitter in COCA, BNC and GloWbE

No.	Word	Word Type	COCA	BNC	GloWbE U.K.	GloWbE U.S.	Total
1.	*twitterverse*	(n.)	38	0	59	84	181
2.	*twitterati*	(n.)	7	0	97	16	120
3.	*twitterstorm*	(n.)	3	0	16	2	18
4.	*twitchfork*	(n.)	0	0	5	0	5
5.	*twinterview*	(n.)	1	0	0	0	1
6.	*twittch*	(n.)	0	0	0	0	0
7.	*sweeple*	(n.)	0	0	0	0	0
8.	*twaffiic*	(n.)	0	0	0	0	0
9.	*adventuritter*	(n.)	0	0	0	0	0
10.	*tweeterbox*	(n.)	0	0	0	0	0

As Table (2) above reveals, the most frequently used term in this group is *twitterverse*, followed closely by *twitterati*, and then (but with much less frequencies) *twittterstorm, twitchfork* and *twinterview*. We may notice that the overall frequencies for *Twitter* are much less than for *weblogs*, most likely because the former was produced in March 2006, whereas the latter have been around since the 1990's. The sections below discuss some patterns found in the types of word formation relating to the words presented in Table (2) as well as their etymology, morphonemics, group systems relating to morphological phenomena, different strata and examples of context in which these terms are found. Examples of context were extracted from GloWbE, excluding those terms which had no frequency. For such terms, the authors provide other examples from the internet.

Portmanteu

Twinterview is a noun which means 'an interview via twitter'. It consists of *twi* (i.e. *Twitter* (n.) - see above discussion) and *interview* (n.), i.e. 'a meeting of people face to face, for consultation' (derived from French). Here, *Twitter* is clipped to *twi*, and shares the same phonological segment of *'i'* with *interview*. Consider the following example from 'Listen to a Londoner' (U.K.):

> *You run a weekly **Twinterview** series on the blog. Which Londoner would you most like to **Twinterview** and why? H & J:; We'd love to **Twinterview** Polly Morgan, a London based artist who specialises in taxidermy (we know, taxidermy again!).*
>
> (http://www.littlelondonobservationist.wordpress.com)

Twitchfork (adj.) *protest is* 'a protest launched on Twitter'. In consists of *twi* (i.e. *Twitter* (n.) and *itch* (from *pitch*, the beginning of the word *pitchfork* 'a large, fork-like tool, which is used on farms for carrying hay but could easily be used as a weapon'. There is an analogy with a 'pitchfork protest' in which ordinary people armed only with the tools of their trade, rise up against their rulers. The word *twitchfork* first began to appear in 2009. By analogy, 21st century protestors reach for computers and mobile phones in order to bombard their targets with verbal weapons. Consider the following example from 'Is McAlpine Going to Sue Me for Saying This?' (U.K):

> *Ms Dent attacks "**Twitchfork** mobs","frothy-mouthed freedom-of-speech defenders" and "publish-and-be-damned" bloggers on the internet. A group of people who incidentally have single-handedly been responsible for the demise of mainstream media.*
>
> (http:www.//inquiringminds.cc)

Twaffic (n.) refers to 'twitter traffic'. *Tw* is a clipping of *twitter* (see above discussion), and *affic* is a clipping of *traffic* which refers to 'the messages or signals transmitted through a communications system', and can be traced back to French, Spanish and Italian**.** Consider the following example from 'Digital Life: Never get Stuck in Twaffic again' (U.K.)**:**

> *Plot your trip the way you would with any sat-nav and* ***Twaffic*** *warns you of any current bottlenecks/roadworks giving you the option to route around them.*
>
> (https://www.independent.ie)

Adventuritter (n.) means 'an adventurous twitterer'. We have the word *adventure,* i.e. an unusual and exciting or daring experience (Latin /old French), and the end of the word *Twitter* following example from 'Need to Know Twitter Speak' (U.K.):

> ***Adventuritter:*** *Slang term used to describe a Twitterer who is adventurous.*
>
> (https://www.kwlms.com**)**

Clippings and Phonemic Overlaps

Twittch (n.) means 'a very, very good friend from Twitter'. Here, we encounter both a phonemic overlap and clipping with *twi* (i.e. *twitter*) and *itch* (i.e. *bitch*). This term is used in the *Twitterverse* in exchange for the word *bitch*. The Germanic term *bicce* originally referred to a 'female dog, wolf, fox, or otter', but was later used to describe an 'unpleasant woman' and is now used like 'guys' (i.e. 'friends'), especially in the case of American English and the growing influence of African American dialects. Consider the following example from 'Urban Dictionary' (U.K.):

> *I love my* ***twittches****!*
>
> (https://www.urbandictionary.com)

Sweeple (n.) means 'sweet twitter people'. It is a phonemic overlap and clipping of *swee* (i.e. *sweet* - 'dear /beloved' (Germanic)) and *ple* (i.e. *people* - 'human beings' (Latin)). Consider the following example from 'Urban Dictionary' (U.K.)

> *My followers are all* ***sweeple.***
>
> (https://www.urbandictionary.com)

Blends with Affixes

Twitterverse (n.) is 'the entire universe of Twitter, including all its users, tweets and cultural conventions'. It consists of *Twitter* and the English suffix *verse* which denotes 'an area of activity or interest or a section of society', originating from the term *universe* (1980s). Consider the following example from 'Michel Hogan - This is not about Brand' (U.K.):

> *I guess this certainly goes to show that you do have to be somewhat careful what you tweet as you never know which ones will stand ignored among the millions per second and which one will be plucked from obscurity and sent careening around the* ***twitterverse.***
>
> (http://www.michelhogan.com)

Twitterati (n.) refers to 'popular users on Twitter; people who usually have large groups of followers and are well known'. It consists of *Twitter,* and the suffix *arti* which is similar to the older *literati,* 'well-educated people who are interested in literature', from the Latin *litter / literatus* 'literature' (Latin) and *letter* 'to form' (English). Consider the following example from 'The Twitter 100: Britain's Titans of the Twittersphere' (U.K.):

> *After taking more than three years to generate its first one billion tweets, it now produces that amount every four days. It has been valued at $10 billion and has spawned its own lexicon with terms such at* ***Twitterati*** *and Twitter storm.*
>
> (http://www.independent.co.uk)

Compounds

Twitterstorm (n.) refers to 'a sudden spike in activity surrounding a certain topic on the Twitter social media site'. *Storm* is of a Germanic origin, meaning 'a tumultuous reaction; an uproar or controversy'. Consider the following example from 'Samantha Brick says Backlash and Bile over Yesterday's Daily Mail' (U.K.):

> *The phone calls were largely from other people in the media -- radio and television researchers -- calling in their droves to ask me to defend myself in the face of the '****Twitterstorm'.***
>
> (http://www.dailymail.co.uk)

Tweeterbox (n.) is 'a twitterer who tweets too much' or 'a person who posts on Twitter, especially on a regular basis'. It consists of *tweeter* and *box*, i.e. 'a container with a flat base and sides' (Greek). This word is actually a play on the compound *chatterbox*, i.e. 'someone who likes to talk too much' (1800's English expression). Consider the following example from 'The Ultimate Guide for Everything Twitter' (U.K.):

> ***Tweeterbox****- Someone who Tweets too much.*
>
> (https://www.webdesignerdepot.com*)*

Facebook

The words in the category of *Facebook* are represented in Table (3), with their frequencies from the three corpora. This is followed by a discussion of their etymology, productivity, frequency (through examination of online corpora), context of use and development. *Facebook* (launched in 2004) is an American for-profit corporation and an online social media and social networking service based in California. Facebook is a compound consisting of the Latin noun *face* (i.e. the front part of a person's head from the forehead to the chin) and the Germanic noun *book* (i.e. a written or printed work).

Table (3). Frequencies of words relating to Facebook in COCA, BNC and GloW

No.	Word	Word Type	COCA	BNC	GloWbE		Total
					U.K.	U.S.	
1.	*facebook*	(v.)	13659	0	37979	49265	100,903
2.	*defriend*	(v.)	5	0	16	20	41
3.	*facebookending*	(v.)	0	0	0	0	0
4.	*sharenting*	(v.)	0	0	0	0	0
5.	*facebookemia*	(n.)	0	0	0	0	0
6.	*facebookeeping*	(v.)	0	0	0	0	0
7.	*facebookee*	(n.)	0	0	0	0	0
8.	*facebookemon*	(n.)	0	0	0	0	0
9.	*facebookering*	(v.)	0	0	0	0	0
10.	*facebooketeer*	(n.)	0	0	0	0	0

As Table (3) above reveals, the most frequently used term in this group is *Facebook*, followed by *defriend*. The rest of the terms in this set do not appear in any of the corpora utilized in this study. This may be explained by the fact that *Facebook* was only introduced in 2004 and the new terminology emerging as a result of this platform is still not very widespread. Examples of context were extracted from GloWbE, excluding those terms which had no frequency. For such terms, the authors provide other examples from the internet.

Blends with Affixes

Facebookending is a verb which means 'in addition to impulsive Facebook check-ins, the act of beginning and ending one's daily computer-related activities with a "quick" Facebook scan for new messages, status updates, and commentary'. The word consists of three parts, namely; 1) *Facebook*; 2) the noun *end*, i.e. 'a final part of something' (OE); and, 3) *ing* is a Germanic suffix 'denoting a verbal action / gerund'. It is a lexical blend since its constituent parts do not represent the meaning of the word as a whole. It may possibly derive from *addend*, i.e. a number which is added to another. Consider the following example from 'Urban Dictionary' (U.K.):

Text: It's 3 am. what's with all the ***facebookending*** *lately?*
(https://www.urbandictionary.com)

Facebook + *ed/ing* as a verb refers to 'the process of uploading photos / contacting friends / looking up contacts (on Facebook)'. It consists of the noun *Facebook*, as well as the traditional suffix endings for past, i.e. *ed*, or present tense, i.e. *ing*. Consider the following example from 'The Cambridge Advanced Learner's Dictionary & Thesaurus (U.K.):

Facebooking *takes up a lot of my time.*
(https:www.//dictionary.cambridge.org)

Defriend (thought to have emerged in the 1800's) as a verb refers to 'removing someone from your network'. The Latin suffix *de* means 'away', and the Germanic noun *friend* means 'a person with whom one has a bond of mutual affection'. Consider the following example from 'Unfriend' or 'Defriend'? Facebook Fans Debate' (U.S.):

> *Another chimed in: "I also use '**defriend**' and have never heard of 'unfriend.' For the record, I'm 24 and use Facebook most regularly. I first signed onto Facebook when I was 19, right when it opened up to my college."*
> (https:www.abcnews.go.com)

Facebookemia is a noun which refers to 'when someone posts on Facebook that they are sick, so when they have called out sick to work, friends have no choice but to accept the fact that the person is sick'. This word is a blend of *Facebook* and *emia*. The suffix '*-emia'* attached to nouns denotes that a substance is present in the blood; (cf. *septicaemia* and *lukemia* (Greek)). Consider the following example from 'Urban Dictionary' (U.K.):

> *worker 1: But Lisa posted on FB that she's "sick as a dog"*
> *worker 2: Yeah,* ***facebookemia*** *again, the only thing she's sick of is working!*
> (https://www.urbandictionary.com)

A *facebookee* is 'someone who is addicted to Facebook'. This is a play on the word *druggy*, i.e. 'someone addicted to drugs'. The Latin suffix *'ee'*, denotes the person affected directly or indirectly by the action (cf. *employee*). The choice of *'ee'* rather than *'y'* may be because *'ee'* tends to denote a person rather than a state; cf. *lucky*, for example. Consider the following example from 'Urban Dictionary' (U.K.):

Check back again, you ***facebookee!***
(https://www.urbandictionary.com)

Facebookering (v.) refers to the process of 'spending ones whole day pointlessly engulfed in the addiction of Facebook'. It consists of *Facebook*, and two suffixes; 1) the Germanic '*er'*,

forming frequentative verbs such as *glimmer, patter*, and '*ing'*. Consider the following example from 'Urban Dictionary' (U.K.):

*I spent yesterday **Facebookering.***

(https://www.urbandictionary.com)

Overlaps

Sharenting (v.) refers to the process of 'sharing every detail of your parenting experience on Facebook' and is an overlap of *share* and *parent,* with the *ing* suffix. *Share,* a Germanic verb, means 'to divide a part or portion of a larger amount'. *Parent* is a Latin word for *mother* or *father*. The Germanic suffix *ing* denotes a verbal action, an instance of this, or its result. The gerund *parenting* also means 'the raising of a child by its parents'. Consider the following example from 'I'm Giving Up Sharenting for the Sake of My Children' (U.K.):

> *According to yesterday's research on "**sharenting"**, taken from Ofcom's annual Communications Market Report, people take on average six photos before choosing a selfie to post online, while young people can take as many as 20 – almost half of those posting selfies say they edit the photo before they share it online.*
>
> (https://www.theguardian.com)

Facebookeeping (v.) refers to 'the process of catching up with your Facebook account after much time of neglect'. There is an overlap, with '*k'* shared by both words. We may make a comparison with *housekeeping,* i.e. to keep, own and manage (old English). Consider the following example from 'Urban Dictionary' (U.K.):

Bryne: Damn, I haven't written back to anyone's post on months
*Melissa: Wow, you have some hardcore **facebookeeping** to do*

(https://www.urbandictionary.com)

Portmanteu

Facebookemon is 'a person who you don't really know, but would like people to think you do'. We may derive this term from *Pokemon* (Japanese, although a loan word from English 'pocket monster'), which refers to a pet monster in a fictional world. Consider the following example from 'Brave New World of Words' (U.K.):

> ***Facebookemon**: The term used for the collection of people, on your Facebook friends list, that you don't actually talk to or know in real life. Related to popular tv programme and game Pokemon, where the aim is to collect as many different Pocket Monsters as possible: Dave: "Hey, my friends collection is getting up to two hundred." Steve: "How many do you really know? Go on, admit it, most of them are **Facebookemon**." John: "Gotta catch 'em all".*
>
> (https://www.iol.co.za/business-report)

Clippings

Facebooketeer (n.) refers to someone who 'knightly stays online on Facebook'. We may derive the end of the word, i.e. the clipping *eteer,* from *muskateer;* a soldier armed with a musket. We have previously discussed *Facebook*, and the Germani suffix *er* denotes a person or thing that

performs a specified action or activity. This word has a humorous element, perhaps drawing on the fictional characters from the novel 'The Three Muskateers'. Consider the following example from 'Urban Dictionary' (U.K.):

> *A "**Facebooketeer**" is a person who wakes up and immediately goes on Facebook & endless "knightly" chat.*
>
> (https://www.urbandictionary.com)

Spelling Changes, Productivity and Lexical Strata Across all Data

Notice that with the words *microblogging*, *liveblogging*, *moblogging*, *blogosphere*, *bloginista and bloggable* there are spelling changes in line with Standard English. *Microblogging*, *liveblogging*, and *moblogging* all double the *'g'* in *blog*, after the addition of '*ing*' following the rule, 'for one-syllable words that end in consonant-vowel-consonant (except *x* and *w*), double the last letter and add -*ing*'. An extra '*g*' is also added to *bloggable*, following the verb-adjective rule, 'if the vowel was short in Middle English (/ɪ ɛ æ ɔ ə/), it uses only a one vowel letter, and is in a syllable ended by a consonant, then the consonant is doubled in spelling, as in *loggable*, *droppable*, *dottable*, *gettable*, etc.

The vowel insertion of '*i'* in between *'g'* and *'n'* in *bloginista* is most probably added to break up the consonant cluster in order to facilitate morphophonemic pronunciation. *Blogosphere* also follows the older word *logosphere*, by adding the vowel *'o'* between the consonants *'g'* and *'s'*, again most likely to facilitate pronunciation and break up the consonant cluster. It seems that whoever coined the word *twittch*, added '*t'* not to be confused with the word *twitch*, i.e. 'give or cause to give a short, sudden jerking or convulsive movement'. However, words in English do not usually contain a double *'t'* before *'ch'* so this is a violation of lexical strata. In relation to *adventuritter*, the *'e'* from *adventurer*, and the *'tw'* from *Twitter* have been omitted, presumably for ease of pronunciation and to contain a similar number of syllables as the word it is based on, i.e. *adventurer*.

With regards to lexical strata, the affixation of *er* in *adventuritter*, is not in violation since it can attach to non-native, as well as native bases. Most affixes native to English attach to both non-native and native bases which explains the acceptability of words such as *twitterati* and *bloginista*. The fact that two native suffixes following each other complies with the lexical strata of English also explains the acceptability of *facebookering*. Some Latin prefixes such as '*de*' and '*un*' are also able to attach to native bases, explaining the acceptability of *defriend*. In general, (with the exception of *twittch*), although the lexical items under investigation are fairly new to the English language, they tend to comply with the lexical strata of Standard English. In terms of productivity, we find that the productive affixes, *'er'* (for a doer of an action), *'ed'* and *'ing'* (for tense) and *'es'* (for plural) are common within the data, such as *facebookering*, *facebooked*, and *twittches*. Yet, as expected, there are no, what might be considered, irregular forms such as the '*en'* suffix, for plural nouns or present perfect tense.

Overall Word Formation

Table (4) below groups together the words over the three sets, i.e. *weblogs*, *Twitter* and *Facebook*, in terms of word formation, type, i.e. whether they are compositional or lexicalized, and overall frequency across the three corpora used in this study.

Table (4). Different Groups of Word Formation, Type and Overall Frequency

Word Formation	No.	Compositional	Lexicalized	Overall frequency	Percentage
Clippings and Splinters	4	*moblogging* *facebooketeer*	*blog* *vlog*	21, 922	59.9%
Blends with affixes	11	*bloginista* *bloggable* *eco-blog* *twitterverse* *twitterati* *facebook/ed/ing* *defriend* *facebookemia*	*facebookending* *facebookee* *facebookering*	11, 245	30.7%
Compounds	6	*microblogging* *liveblogging* *blogosphere* *mummy blog* *twitterstorm*	*tweeterbox*	3, 383	9.2%
Portmanteau	4	*twinterview* *adventuritter*	*facebookemon* *twitchfork*	6	0.2%
Clippings and Phonemic overlap	5	*facebookeeping* *twaffic*	*twittch* *sweeple* *sharenting*	0	0%
Acronyms	1	*GRWM*		0	0%
Initialisms	1	*OOTD*		0	0%
Total	**32**	**21**	**11**	**36,556**	**100%**

According to the data, we can see from Table (4) above that clippings and splinters have the largest overall frequency across the three corpora (59.9%), followed by blends (30.7%) and then, compounds (9.2%). There are also a larger number of compositional forms than lexicalized. Blends with affixes are also the largest group in terms of type of word formation, followed by compounds. Out of the 32 lexical items under scrutiny, only 13 had some level of frequency across the three corpora. Table (5) below shows these frequencies. Consequently, 17 of the lexical items had no frequency whatsoever across all three corpora.

Table (5). Highest Frequencies of Lexical Items across COCA, BNC and GloWbE.

No.	Word	Word Type	COCA	BNC	GloWbE		Total
					U.K.	U.S.	
1.	*facebook*	(v.)	13659	0	37979	49265	100,903
2.	*blog*	(n. / v.)	5,968	0	7047	8707	21,722
3.	*blogosphere*	(n.)	278	0	986	1567	2831
4.	*microblogging*	(n. /v.)	50	0	244	143	437
5.	*vlog*	(n. /v.)	9	0	78	107	194
6.	*twitterverse*	(n.)	38	0	59	84	181
7.	*twitterati*	(n.)	7	0	97	16	120
8.	*liveblogging*	(n. /v.)	2	0	76	31	109
9.	*defriend*	(v.)	5	0	16	20	41
10.	*twitterstorm*	(n.)	3	0	16	2	18
11.	*moblogging*	(n. /v.)	1	0	3	2	6
12.	*twitchfork (protest)*	(n.)	0	0	5	0	5
13.	*twinterview*	(n.)	1	0	0	0	1

As is evident in Table (5), the most frequently used word is *Facebook* (as a noun or verb), and hence we may conclude that it is the most productive according to the corpora employed. This is followed by *blog*, and then to a much lesser extent, *blogosphere*, *microblogging* and *vlog*. The ever-growing popularity of Facebook with up to date, 2 billion users globally, may explain this phenomenon. The fact that *blogs* are also extremely popular and have wide ranging applicability with an estimated 173 million blogs online, may also explain this word's high frequency.

Limitations and Further Research

The authors recognize that the results can only be used indicatively in light of the sample number, and do not attempt to make conclusive generalisations based on these findings. Further research may include a much larger number of words, as well as various social media platforms, such as *YouTube, Snapchat, Flicker*, *Instagram* and many others. Other corpora could also be incorporated to assess the frequencies of these lexical items across more extensive genres, as well as other varieties of English other than British and American English.

Conclusions

In the present study, both quantitative and qualitative methodologies were employed to provide a comprehensive understanding of the findings and the quantitative results generated by the three corpora. The qualitative analysis focused on examining the lexical items under investigation in terms of their metaphorical meanings, their relevance to social media platforms, and their derivation from earlier forms of the English language. This qualitative approach further enabled an assessment of the extent to which these lexical items conform to, or deviate from, the lexical strata of Standard English. The analysis revealed that, with the exception of the word '*twittch'*, the novel lexical items examined in this study conform to the lexical strata of Standard English.

The quantitative component of the study involved assessing the frequency of the lexical items under investigation and, by extension, their productivity. In many cases, no occurrences of these items were found across the three corpora, reflecting both their relative novelty and the extent to which they have yet to be fully integrated into spoken and written English, despite their prevalence on the internet. This interpretation, however, requires further verification through systematic research, given the dataset and corpora provide only a partial picture of social media–related word formation, making wider generalizations difficult. The findings also indicate that numerous social media–related lexical items are not yet widely attested, particularly in written texts, as evidenced by their absence across three corpora considered representative of British and American English.

References

- Algeo, J. (1977). Blends, a structural and systemic view. *American Speech*, *52*(1/2), 47-64.
- British National Corpus (BYU-BNC) - BYU corpora. https://www.corpus.byu.edu/bnc
- Cacciari, C. & Padovani, R. (2012). The development of figurative language. In Spivey, Michael J., Ken Mcrae and Marc F. Joanisse. (Eds.), *The Cambridge handbook of psycholinguistics.* (pp.505-522). New York: CUP.
- Caruntu-Caraman, L. (2022). The Productivity of lexemes of English origin. *Philologia*, *(3)*2, 32-41.
- Corpus of Contemporary American English (COCA). *COCA*. https://corpus.byu.edu/coca
- Corpus of Web-Based Global English. *GloWbE*. https://www.corpus.byu.edu/glowbe
- Debjani, R. (2014). Cinema in the age of digital revolution. *International Journal of Interdisciplinary and Multidisciplinary Studies*, *1*(4), 107-111.

- Emayakre, N.N., Kuufaar, C. and Gimode, J.K. (2022). Magnitude of productivity: Inflectional affix or derivational affix. *Journal of Arts and Social Sciences.4*(1), 1-17.
- Fernández-Domínguez, J. (2010). Productivity vs. lexicalization: Frequency-based hypotheses on word-formation. *Poznan Studies in Contemporary Linguistics. 46*(2):193-219
- Gries, S. T. (2004). Shouldn't it be breakfunch? A quantitative analysis of blend structure in English. *Linguistics*, *42*(3), 639–667.
- Harald, B. & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics 29,* 801-844.
- Haspelmath, M. (2024). Inflection and derivation as traditional comparative concept. *Linguistics*, *62*(1), 43–77.
- Haspelmath, M. & Simms, A.D. (2010). *Understanding morphology*. (2nd ed.). London: Hodder Education.
- Hiskey, D. (2012). *The difference between an acronym and an initialism.* https://www.todayifoundout.com
- Kosur, H. M. (2012). *Word formation: Coinages, nonce words, borrowing, and calquing.* https://www.brighthubeducation.com
- Lieber, R. (2009). *Introducing morphology*. Cambridge: CUP.
- Schoenherr, S.E. (2004). *The digital revolution*. http://www.history.sandiego.edu