English Lexical Enrichment: Methods and their Frequency/Productivity in 1785, 1885, and 1985

Rozgar Y Omar¹, Nechirvan Hassan Jawzal²
¹Department of Mathematics, College of Science, Salahaddin University-Erbil
²Department of English Language, Faculty of Humanities, University of Zakho

ABSTRACT
There are a number of different categories of enriching English vocabulary, which are grouped into smaller methods based on the way they function. Some methods are productive, while some others are creative. Besides, some methods are more or less frequent/productive than others which may be due to a number of reasons such as the nature of the language itself and possibly the period the language is undergone on. Therefore, three different years from different centuries, precisely 1785, 1885, and 1985, are taken from Oxford English Dictionary (OED) to discover the frequency/productivity of the methods in each. The concrete results show that the language and time play a role in the frequency and productivity of each method. They also indicate that the frequency of some methods has gradually increased while of some others decreased.

Keywords: lexical enrichment, word-formation, frequency, productivity, OED.

1. Introduction
There are different reasons and distinctive ways to enrich English vocabulary. According to Minkova and Stockwell (2009: 5-22), there are two main concepts of making new word meaning and neologism, namely regular word-formation and new word-creation. The first one is productive while the second is creative.

Regular word-formation includes the processes of attaching affixes to a base or a stem (affixation), converting the syntactic category of a word without changing its form (conversion), and connecting words/combining forms (compounding). The second process, on the other hand, which is new word-creation, is recognizable by its individual usage. It encompasses different procedures of generating new words such as creation de novo, blending, clipping, back-formation, abbreviation, eponymy, and other sources like onomatopoeia and reduplication.

Two other methods of vocabulary enrichment, which is termed ‘foreign’ because of the way they look to the native language, are borrowing and folk-etymology. Borrowing is the process in which a language, instead of assisting from its lexicon, loans vocabulary from another language (Katamba, 1994:133) without a major change, while, folk-etymology is a false assumption of a form ( RATIH and GUSDIAN 2018: 27) and modeling it “under the influence of some other word” (McAtee, 1951: 90).

Bauer (2001) examines different methods to find an answer for whether a process of coining a new word is productive or creative. Then he finally concludes that it is very problematic to distinguish between productivity and creativity.

Nonetheless, instead of arguing to respond to his claim, in this work, the theory of Minkova and Stockwell is taken into account that some processes can be productive and some others can be creative. So, although there are different methods of lexical enrichment in English and many studies have been conducted about it (Plag 2018, Štekauer and Lieber 2006, Bauer 1983, Kiparsky 1982), the objective is to examine three different years from different centuries to discover which processes appear to be more or less frequent/productive. In other words, the aim of this work is to find whether centuries can play the key role in having the lexical enrichment categories being more
or less frequent. The significance of the research lies in the data collected to determine that centuries can affect the frequency/productivity of the categories. Due to the word limit and the size of the current research, this research couldn’t cover the data from the three centuries as a whole as it would require years of academic research.

The online Oxford English Dictionary (OED) will be used to find the data. Then the results in the light of a possible discussion will be coped with. It must be stated, though, that the entities found on OED at a specific time does not mean that the word is created just then. To a word to enter into a dictionary, it goes into a language change process until it will be recorded as an entity in that dictionary. Language starts to change once an individual or a group of individuals create a neologism (i.e. a first-time new word occurrence) which by time is adopted by more speakers than by a community, and then it is finally accepted as a standard usage by the public (McMahon, 1994:8-9). So the words found in our data might have been created sooner than the OED shows, but, because there cannot be any record to trace, the entities we find will be considered as first time used/created.

2. Literature Review

As discussed above, there are different methods of lexical enrichment such as regular word-formation, new word-creation, and foreign enrichment. Each is dealt with separately in this section.

The first category of regular word-formation is **affixation**. It is a process to produce a new word by adding a listeme (prefix, suffix, interfix or infix) to an existing lexeme (stem/root/base). Due to their rareness and restriction, neither interfix nor infix will be studied here.

A **prefix** is an affix attached to the beginning of a lexeme (Tokar, 2012: 56). Semantically, prefixes can be categorised into different groups (Plag, 2018: 98) such as (a) quantifying prefixes as uni-, which quantifies the meaning of the lexeme and expresses ‘one’ in unicolour, (b) negation prefixes as in- in infertile, (c) locational prefixes as inter- in interstation, (d) temporal prefixes as pre- in prehistoric, (e) prefixes of degree as sub- in substation, and (f) reversative/privative prefixes as de- in deerate (Montero-Fleta, 2013: 273).

Also, **suffixes**, attached to the end of an element, can be classified into different groups such as (a) nominal suffixes, where they will be added to root/stem and the syntactic category, if it is not already a nominal, will change into a nominal phrase, such nominal suffixes are -ance in acceptance, -ee in employee, -hood in falsehood. Other categories of suffixes are (b) verbal suffixes which derive verbs mostly from nouns and adjectives, such as -fy in beautify and -ise in familiarise; and (c) adjectival suffixes -able in employable and -ive in active. The last group is called (d) adverbial suffixes as in -ly in quickly and -ward in homeward.

The second category is **conversion**, which is a method of converting a word class of a word into another. For instance, the meaning of to father is derived from the noun a father (Harley, 2006:105). Sometimes it is also called ‘zero derivation’ because the shape and the form of the ‘input word’ will not be altered. According to Katamba (1994: 48-49), there is one possible approach to distinguish what syntactic category converted words are which is only by the grammatical context they are in.

The last method of regular word-formation is **compounding**. A compound is the combination of more than one free morpheme (words/combing forms) sharing a conceptual relation. On the other hand, the second category of lexical enrichment is new word-creation. It is subcategorised into different methods.

**Creation de novo** is a process of introducing a new word which happens unexpectedly, rarely to create a word totally from scratch. The process is sometimes
called word-manufacture (Bauer, 2006: 498) because often they appear as trade-marks, such words are Exxon, dongle, and Kodak.

**Blending** refers to creating a new lexeme by clipping two different words then blending them to stand as one morpheme, often by clipping the last letters of the first morpheme and clipping the first parts of the second, where Kelly (1998: 579) believes components “fall at major phonological joints”. Chillax (chill + relax), infomercial (information + commercial), shortle (chuckle + snort) are examples of blending.

**Clipping** is another creative process of coining a word by clipping off or shortening an existing, usually multisyllabic, word. Possibly, due to its length and difficulty of pronunciation, speakers clip common words (Harley, 2006: 95), like gymnasium to gym, Internet to net, vegetarian to vegan, and jammies from pyjamas.

Another way of creatively introducing new words is **back-formation**. It is a process when a “presumed affix” (usually a suffix) is clipped off from an existing word. For instance, the word sedate is derived from the word sedative (Brinton, 2000: 98). It came to use when the speaker considered the recognizable -ive at the end of the word is transparent. Other examples of back-formation can be seen in televise (from television), orate (from oration) and burgle (from burglar).

Another shortening mode of making new words is **abbreviation**. This is a process, whereby a long compound is clipped off to one word, to take the initial sound of words and combine them into one element (Grange and Bloom, 2000: 2). It is phonologically subcategorised into two types, namely acronyms and initialisms (or alphabetism). The difference between both is when the set is pronounced as a single component it is an **acronym** as gif (graphic interchange format) pronounced as /ɡɪf/ (or /dʒɪf/). Yet, the abbreviated KRG word is an example of **initialisms** because each letter of KRG, which stands for 'Kurdistan Regional Government', is pronounced individually, /keɪ ɑː r dʒiː/.

**Eponymy**, which is a creative process of making new words, is based on the derivation of names. A great number of discoverers in medicine, biology and physics are examples of eponymy (Minkova and Stockwell, 2009: 19). It can be classified into different groups according to personal names (quisling, from Vidkun Quisling), geographical names (jean, from Italian city Genoa), mythical characters (gorgon, from Greek Mythology), and commercial brand names (Netflix as in “Let’s Netflix and chill”).

Other methods of new word-creation, dependant on the pronunciation of the lexeme, are onomatopoeia (or echoing/imitative) and reduplication. **Onomatopoeic** words are those words that for listeners can be readily understandable from the sound of the word, as crack, vroom, his, and beep, where these sound like a noise.

**Reduplication**, also based on the pronunciation of the word, is a process to make a compound noun by repeating a word or part of a word (Nadarajan, 2006: 39), and based on juxtaposition, it can be classified into two categories: (a) Full: night-night, ack-ack, and gee-gee, and (b) Part: tell-tale, funny-bunny, on-off, eff-off, muck-tub, slip-slop, even-steven, and man-o-man (Rastall, 2004: 40).

Lastly, the foreign methods of lexical enrichment are **borrowing** and folk-etymology. A language borrows a word from another for different reasons, such as culture (coffee), politics (parliament), religion (hajj), economy (market), and prestige (beef). As defined above, **folk-etymology** is a replacement of another form, and can happen on two levels: functional (word-formation change) and lexical change (in terms of phonetics/semantics) (Coleman and Kay, 2000: 22, Michel, 2015: 1) such as cockroach, bridegroom, bus, -gate, and mayday (Moore, 2014: 18-20).
3. Methodology

As shown above, there are different processes of lexical enrichment in English. Some are more productive and more frequently used, whereas some others are rarely used. Also, a large number of studies in this field have been conducted such as in Cutler (1980), Štekauer (1998), Dressler and Ladányi (2000), Bauer (2001), Plag (2006), Fernández-Domínguez (2010), and Hacken and Panocová (2013). However, few of them concentrated on frequency/productivity of each process in a certain period comparing to another period. Therefore, this study will compare each category in three different periods of time specifically 1785, 1885, and 1985 so as to determine which method is overall frequent/productive. These specific years, without any linguistic significance, are selected randomly from different centuries because taking data from each century as a whole would need years of work and cannot fit into a research study due to word limit. The main idea here is to find how English enriched its vocabulary in different centuries, at particular points. Our data are from the online OED (Oxford English Dictionary) which is described as a ‘giant source’ (Good, 2012). The process of collecting data primarily started by Advanced Searching of words coined in 1785, 1885, and 1985.

4. Results and Discussion

According to the collected data, 370 words have enriched English vocabulary in 1785, 953 in 1885, and 187 in 1985. However, some of the words in each period (34 in 1785, 148 in 1885, and 26 in 1985) were not counted into the data for different reasons which are discussed below.

a) Unknown/uncertain origin: bawd, donkey, fuji, natty, olivart, and tosher.

b) Variant/altermation/obsolete form: caneva, geezer, hackit, knobble, pisette, and squitch.

c) Change in meaning: buzz (as another verb with different semantic meaning), ducker, nickle, and sledger.

d) Mixed methods: buckminsterfullerene (eponym + suffix), conceptualist (concept + ualist or directly from the French conceptualiste), interhyal (inter- + hy-(hyoid) + -al), and nanny (nan + y or Nanny pet form names for Anne/Agnes).

e) Unclear method: autogravure (OED states that it is a loan word from French combined with an English element, though auto- and gravure had been used in 1876 and photogravure in 1873. Therefore, it is either compound word or a blended one), Brown Bess (even though OED states that it is a compound of Brown (adjective) + female name Bess, Grose (1931: 54) gives an explanation and says it is “more probably a personal name”), barrack (OED is not certain whether it is originally from Australian (alteration of borak) or an entity from English Dialectical Dictionary which refers to the word as northern Irish (Wright 1898: 171), plus there has been barrack in 1686 with different meaning. Though, still, the method through which this word is formed is not determined), Locomobility (OED states that it is a compound from loco- + mobility, while first use of loco is registered in 1786, which means the word is not formed by compounding. Also, the words locomotive and locomotion appeared first in 1612 and...
1646 respectively. Therefore, it can be either blending or borrowing from French *locomobilite* (appeared in 1777), *pitmatic* (OED says it is a compound word, but it is derived from *pit* + -matic (of mathematic) and there has been *pitmatics* in 1844, so it is either blending or back-formation), and *spatch-cock* (OED suggests that it is an “abbreviation of a *dispatch cock*”, while it has been stated by Palmer (1882: 366) that the word came from “*spithcock*”. In Sayers (2012: 4-5), it is claimed that the origin of the word is still unclear). Hence, the above mentioned entities are not counted in our data.

Relevant to OED’s etymology, dozens of entities are not categorised or are miscategorised into different enrichment methods, such as *chug*, *corella*, *electret*, *Mesmerite*, *pulvinarian*, and *skatepunk*. OED claims *chug* (v.) is an onomatopoeia but it is conversion from *chug* (n.) used in 1866. For *corella*, OED does not give any information about its etymology, while Merriam Webster Online states it is a “native name in Australia”, which makes it eponym. However, Dixon et al. (2006: 83) state that it is a borrowed word from Wiradhuri, probably from *garila*. Based on OED, *electret* is a compound and it is formed from *electr*-(in electric) + -et (in magnet); hence it is a blend. *Mesmerite* is claimed by OED to belong to double-formation (a proper name *Mesmer* combined with -ite (eponymy + suffix)). However, it would be true if *mesmerism* was not formed in 1784. Therefore, *mesmerite* can be observed as a base + suffix, which is counted as suffixation. *Pulvinarian* is suggested to be a loan in OED while its base, *pulvinar*, was first used in 1599, so it is formed through suffixation. For *skatepunk*, OED says it is a compound word, but it is a blend, not from *skate* + *punk* but *skate*(board) + *punk*. Thus, all the data are calculated and categorised based on, not OED’s etymology but, our knowledge to lexical enrichment.

In 1785, 336 entities have been included. As Table 1 illustrates, in total, borrowing appeared to be the most frequent method with 114 words (33.9%). The words are borrowed from these languages: Afrikaans (*skoff*), Arabic-Persian (*burkundaz*), Cree (*wapacuthu*), Dutch (*veld*), French (*spicule*), German (*muggent*), Hindi (*zilah*), Italian (*studio*), Latin (*cembra*), and Sanskrit (*yoga*).

**Table 1: methods of lexical enrichment in English in 1785, by number and in percentage**

<table>
<thead>
<tr>
<th>Methods</th>
<th>Prefix</th>
<th>Suffix</th>
<th>Conversion</th>
<th>Coexisting</th>
<th>Blending</th>
<th>Clipping</th>
<th>Base-formation</th>
<th>Abbreviation</th>
<th>Epithet</th>
<th>Onomatopoeia</th>
<th>Reproduction</th>
<th>Borrowing</th>
<th>Cretica de loco</th>
<th>Folk-etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>35</td>
<td>87</td>
<td>16</td>
<td>70</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>114</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>%</strong></td>
<td>10.4</td>
<td>25.9</td>
<td>4.8</td>
<td>20.8</td>
<td>0.9</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>2.1</td>
<td>0.6</td>
<td>35.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The second most productive method is suffixation by forming new words through suffixes such as *badgering*, *balloonery*, *centrifugally*, *cocombality*, *cryptogamic*, *hulky*, *imperialize*, *jokish*, *neologist*, *quizzical*, *reproductive*, *revolucy*, *servitorship*, *spoutless*, *theorism*, and *wigsly*.

The next method that English lexicon based on is compounding 70 entities, combining a word to another (as *brushman*, *nipcheese*, and *sea-bath*), a combining form to a word (as *grandpa*), a word to a combining form (as *balloonomania*), and a combining form to another (as *novemfid*).

Another approach to word formation is prefixation, with 35 entities, such as *belittle*, *bitterate*, *disintegrity*, *inattentive*, *obulate*, *re-edit*, *semi-palmate*, *subvariety*, and *unimpeachable*. In addition, conversion happened 16 times from noun to verb (as *doss*, *hump*, and *prong*), verb to noun (as *gamble*, *rut*, and *spank*), verb to adjective (as *emarginated*), and noun to adjective (as *nits*).

For the lesser frequent/productive methods, there were 7 eponyms (such as *Constantia*), 3 clipped words (as *chit* from *chitty*), 2 onomatopoeic words (as *flab*), 1 back-formation (*rust* from *rusty*), and 1 reduplication (*goody-goody*). However, there was no record for the other categories.

**Table 2: methods of lexical enrichment in English in 1885, by number and in percentage**
As noted above, a bigger number of coinages enriched English lexicon in 1885, which are 805 apart from the uncounted ones. As seen in Table 2, the most frequently used method is suffixation with 287 entities (35.7%), such as acetylate, algebraize, Americanesque, baglet, bankster, borrowee, cameldom, champignon, disburseable, geomalism, horniness, maltate, maturish, specificize, spiderling, tabulator, and weirdsome.

By loaning from more than 30 languages, borrowing comes second with 209 entities, such as adenine (German), bandar (Hindi), bigos (Polish), charoset (Hebrew), conus (Latin), coperta (Italian), corella (Wiradhuri), dyslexia (French), hale (Hawaiian), hijab (Arabic), ilmoran (Masai), jnani (Sanskrit), Kamba (Bantu), kaza (Turkish), lactoglobulin (Danish), lakatoi (Papuan), levada (Portuguese), mandor (Malay), Narodnik (Russian), Nestilingmiut (Western Canadian Inuit), Nimpkish (Kwakiutl), parakelya (Guyani), pasillo (Spanish), stolkjaerre (Norwegian), taclobo (a language of Philippines), tagma (Greek), Tamashek (Berber), tansu (Japanese), Taranchi (Jagatai), Tet (Vietnamese), ushabti (Egyptian), and witloof (Dutch).

The other method English enriched its lexicon is compounding, with 141 entities, by combining two words (such as bustline and cocksucker), a combining form and a word (as acetocarnine and hydrotherapeutic), a word and a combining form (as brickscape and dichoptic), and two combining forms (as aerophilous and zoophile). Relatively, another productive method was prefixation, attaching 76 prefixes to already existed words such as amphi- (in amphiaster), an- (anaerobiotic), circum- (circumcircle), de- (deprint), dys- (dysmenorrhagia), epi- (epicuticle), hemi- (hemicardiate), in- (inassimilation), mis- (miskick), off- (offprint), out- (outscore), post- (postjudiced), pro- (pragamous), re- (rebid), sub- (subculture), under- (undertint), and up- (upwell).

In 1885, there has been a conversion with 34 entities, such as bike, nursery, and telepath (noun to verb), balter, leave, and replot (verb to noun), housebroken and jack (verb to adjective), made-to-measure (noun to adjective), and multiplex (adjective to verb).

Also, English enriched its lexicon with 24 eponyms (such as Addisonian, Bismarck, Haarlen, Mason, and Pelton), 15 clipped words (as agar, crypto, and talkee), 5 blended words (as electret, Insinuendo, and museology), 5 onomatopoeias (as hallali, pleep, and yeow), 4 back-formed words (as crust-hunt and mantle-make), 3 abbreviated ones (as OUDS, Nd, and Pr), 1 reduplication (yum-yum), and 1 folk-etymology (berceanette).

Comparing to both previously mentioned years, 1985 has made smaller growth to the English lexicon with 187 entities. By excluding 26 of them for the reasons mentioned above, 56 entities (34.8%) belonged to compounding, such as offshore (word + word), multibuy (combining form + word), footballene (word + combining form), and ovice (combining form + combining form). Suffixation comes next with 34 entities, such as conspiracism, garmento, geekdom, malness, metabolan, persher, presidiential, Ramboesque, stripped, tanky, and tungeon.

### Table 3: methods of lexical enrichment in English in 1985, by number and in percentage

<table>
<thead>
<tr>
<th>Method</th>
<th>Prefix</th>
<th>Suffix</th>
<th>Conversion</th>
<th>Compounding</th>
<th>Blending</th>
<th>Clipping</th>
<th>Back-formation</th>
<th>Abbreviation</th>
<th>Eponymy</th>
<th>Onomatopoeia</th>
<th>Reduplication</th>
<th>Replotting</th>
<th>Creation de novo</th>
<th>Folk-etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1985</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>%</strong></td>
<td>3.4</td>
<td>2.1</td>
<td>9.6</td>
<td>2.2</td>
<td>5.9</td>
<td>0.6</td>
<td>25.4</td>
<td>24.9</td>
<td>11.8</td>
<td>11.8</td>
<td>7.3</td>
<td>7.3</td>
<td>16.1</td>
<td>3.4</td>
</tr>
<tr>
<td>h/n</td>
<td>14</td>
<td>8</td>
<td>14</td>
<td>7</td>
<td>8</td>
<td>14</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

Dissimilar to 1785 and 1885, blending appears to be frequent with 19 entities, which makes 11.8% of the total lexical enrichment in 1985, as illustrated in Table
3. The blended examples are *an-dram*, *boxercise*, *hoolivan*, *Japanimation*, *monery*, *rapso*, *snarfle*, and *tankini*. Borrowing appears to be less frequent than the previously mentioned years with 16 entities. They are borrowed from Afrikaans (*potjie*), Arabic (*intifada*), French (*Négrette*), Hebrew (*Krav Maga*), Italian (*guido*), Japanese (*kaizen*), Latin (*oviraptorosaur*), south American Spanish (*inti*), and Vietnamese (*banh mi*). Conversion seems to have occurred 11 times only from noun to verb (such as *focus-group* and *snowboard*) and verb to noun (as *crash-and-burn* and *mosh*). Comparing to the other two years, the use of the clipping method increased and made 5% of the 1985 lexicon with 8 words, such as *dexter* (from *poin-dexter*), *hox* (from *homeobox*), and *sim* (from *simulation*). With closer number, prefixation comes next with 7 entities as *off-guard*, *overramped*, *post-truth*, and *unattributively*. Abbreviation occurred 4 times as *BOGOF* and *WIMP*, so did eponymy as *benji* and *Rambo*. There was 1 creation de novo for *Semtex*, and so was back-formation for *electroporate*.

By comparing the three years in percentages, Figure 2 shows that suffixation and compounding played a significant role in lexical enrichment in English. Suffixation introduced 25.9% in 1785, 35.7% in 1885, and 21.1% in 1985, compounding made 20.8%, 17.5%, and 34.8% in 1785, 1885, and 1985 respectively.

<table>
<thead>
<tr>
<th>Year</th>
<th>Suffixation</th>
<th>Compounding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1785</td>
<td>25.9%</td>
<td>20.8%</td>
</tr>
<tr>
<td>1885</td>
<td>35.7%</td>
<td>17.5%</td>
</tr>
<tr>
<td>1985</td>
<td>21.1%</td>
<td>34.8%</td>
</tr>
</tbody>
</table>

Figure 2: methods of lexical enrichment in English in 1785, 1885, and 1985 in percentage

Some methods appeared to be frequent/productive but their percentages declined from 1785 to 1985 such as prefixation and borrowing. Prefixation dropped from 10.4% to 4.3%, and borrowing from 33.9% to 9.9%. Contrariwise, methods such as blending and clipping inclined, from 0.0% to 11.8% and 0.9% to 5.0 % respectively. The remained methods appeared to be evenly less in their frequency/productivity.

5. Conclusions

The aim of this study was to analyse different methods of lexical enrichment and their frequency/productivity in English in different periods namely 1785, 1885, and 1985. In light of morphology, three different categories have been shown: regular word-formation, new word-creation, and foreign. Then each category has been classified into different methods.

After having collected data for each year based on OED, there are some words whose etymologies are not clear to trace, introduced by mixed methods, and some others are miscategorised by OED. Therefore, in light of the given argument, the data are calculated based on etymologically linguistic evidence.

It has been found that the methods vary in enriching English vocabulary. While, suffixes and compounding are productive in all the three years, some other methods are infrequently used such as back-formation, eponymy, onomatopoeia, reduplication, creation de novo, and folk-etymology. The rest seem to decline as prefixes and borrowing or to incline such as blending and clipping.

6. References

27. Palmer, A. S. (1882). Folk-etymology: A dictionary of verbal corruptions or words perverted in form or meaning, by false derivation or mistaken analogy. G. Bell and sons.