International Journal of Arts and Social Science www.ijassjournal.com
ISSN: 2581-7922,
Volume 7 Issue 12, December 2024

Measuring Morphological Productivity of Suffixes Used in Verb-Based Nouns A Corpus-Based Study

Adil H. Jouda, (M.A.)
English language and linguistics University of Baghdad College of Arts Department of English

Abstract: Language growth and flexibility of use are related to how productive the processes of word formation are, particularly the process of derivation manifested by the use of many suffixes. As productivity is a gradable property, i.e., not uniform, suffixes vary in their productivity.

This study aims at measuring the productivity of suffixes used in verb-based nouns. Bauer's model (2001) has been adopted to achieve that goal. Nineteen suffixes have been checked in a corpus of (4456) words in the British National Corpus by using a web-based system, i.e., the Sketch Engine programme .

The study concludes that the productivity of suffixes used in verb-based nouns is relative to their frequencies of use. The most highly used suffix is – ation standing for one-third of the whole corpus and the second most highly utilized one is -ing, representing one-fifth.

On the other hand, the non-productive suffixes seem to have a very low level of frequency in the whole corpus indicating that these suffixes have been fossilized and no more functioning in the creation of new words to address novel concepts, i.e., they have become archaic. This case is represented by the suffixes –t and –th, where other alternative suffixes have come into play. Thus, this study has shown that suffixes exhibit different levels of productivity and consequently different rates of contribution in the adaptability of the language to the new concepts and in the flexibility of use.

Keywords: productivity, derivation, suffixes, word formation, corpus analysis

I. The concept of productivity

One of the properties of human language is productivity. According to Crystal (2003, p. 374), productivity is defined as "a general term in linguistics to refer to the creative capacity of language users to produce and understand an indefinitely large number of sentences." It is also used to refer to the creativity of human languages which use finite means to produce an infinite number of words or utterances (Katamba, 2006:74). Brown and Miller (1980, p. 224) define productivity by highlighting the general applicability of a rule that "the rules are productive insofar as they apply to a range of stems". The term is also used with a rather restricted sense referring to certain language features or patterns that the morphological patterns are said to be productive if they can be systematically extended (Boij, 2005:18). For instance, the derivation of nouns ending in –er from verbs is productive in English, but deriving nouns ending with –th from adjectives is not. Bauer (2001, p. 25) states several definitions quoted from Rainer(1987) with qualitative and quantitative foci that productivity is defined in terms of frequency of the output words, or in terms of the available bases, i.e., the input category or even the possibility of forming new words.

II. Language productivity

One of the design features that distinguishes human language from other non-human animal communication systems is productivity. According to Crystal (2010, p. 13), productivity means that "the potential number of utterances in any human language is infinite". Katamba (2006, p. 74) also defines it as "the capacity of all human languages to use finite means to produce an infinite number of words or utterances". Trask (1999, p. 5), referring to the uniqueness of human language, mentions productivity or open-endedness as the ability to use language to produce novel utterances that we have never said or heard before (Aitcheson, 1992, p. 15).

III. Process productivity

Talking about the English word-stock, it is quite noticeable that it is getting richer and fuller, reflected in the increasing number of words almost quite regularly. Favor is due to the processes of word formation employed by users of English. The constant evolution of new words and new uses of old words can be viewed as a sign of vitality and creativeness as the language is shaped by the needs of its users (Yule, 2010, p. 53).

The processes listed in morphology and linguistics books (Stageberg, 1981, pp. 120-127& Yule, 2010, pp. 53-59) include, but are not limited to, coinage, borrowing, blending, clipping, back-formation, compounding and reduplication. As far as productivity is concerned, as a gradable property of morphological rules, some of these processes are said to be more or less productive than others (Shopen, 1985, p. 16). Compounding and derivation are very productive sources of new terms in English (Aitcheson, 2012, p. 198&Yule, 2010, p. 55).

Compounding is the joining of two or more words into a single word, e.g. high school, butterfly. Compound words whose meanings can be perceived from the combination of the two constituents are called semantically transparent such as the nouns (workman, boyfriend), the adjectives (good-looking, low-paid), and even verbs such as (downpour, double-click). The other type of compound words whose meanings cannot be recognized from the combination of the two constituents are known as semantically opaque, such as nouns (hotdog, in-laws) verbs (pass away, call-off), and adjectives (devil-may-care)(Stageberg, 1981, p. 121, Gleason and Ratner, 1998, p. 166&Haspelmath, 2002, p. 85).

Derivation, on the other hand, is accomplished by a large number of "bits" of the English language. These small bits are known as affixes. The affixes are classified into prefixes attached to the beginning of the words (e.g., mis-, dis-, un… etc.) or suffixes appended to the end of words such as (-ation -less, -ish, -ing, -ed… etc.)(Stageberg, 1981, p. 89&Yule, 2010, p. 59). Aitcheson (2012, p. 198) states that affixation is the commonest way of forming new words in English and that some affixes are enormously productive such as the affix –ness, which is primarily attached to adjectives, like *happy* and *good*, to produce nouns as in happiness and goodnessand it can also be attached to phrases as in self-centeredness.

IV. Morpheme productivity

According to Gaskell (2007, p. 175), the psycholinguistic description of human communication should explain how language inputs and outputs are structured in order to produce or comprehend the speaker's intended meaning. Meaning is conveyed through words and morphemes which are assembled together in numerous ways to convey the message. The morpheme is seen as the "smallest functioning unit in the composition of words (Crystal, 2003, p. 300). It is also defined as the minimal meaning-bearing unit in the language or the minimal distinctive unit of grammar (Crystal, 2003, p. 300). Accordingly, a distinction can be made between lexical and grammatical morphemes. The lexical morphemes are used to build new words and they carry the messages that we convey as in the compound words e.g., blackboard, and suffixes like –ize, -ship, -ment. The grammatical ones, on the other hand, express grammatical relationships between the word and its context, such as -s plural and –ed past tense morphemes (Yule, 2010, p. 68). As building blocks of words, morphemes are classified into derivational and inflectional based on certain characteristics for each type (Stageberg, 198, pp. 92-95) (Gleason, 1998, p.164) (Fernandez and Cairns, 2010, p. 194).

As for the productivity of morphemes, Crystal (2010, p. 374) indicates that there are three types, i.e., productive, non-productive, semi-productive. The morpheme is said to be productive if it is repeatedly used to produce other instances of the same grammatical category e.g., using –ed past tense where any regular verb in the present tense can be made past tense by the attachment of this morpheme (Roach, 2009, p. 83). Non-productive morphemes do not have such potential, for instance the irregular noun mouse- mice, or the irregular verb speak-spoke, or even the irregular adjective good- better, as these instances cannot be utilized to produce further plurals, past tense forms or comparative adjectives respectively. The third type is semi-productive morphemes that have limited or occasional creativity such as the negative prefix un- which is sometimes used to form the opposites, e.g., *happy –unhappy*, but not *sad –*unsad*.

Fernandez and Cairns (2010, p. 194) refer to the storage and retrieval of bound morphemes indicating that words made by affixing productive morphemes are not stored lexically, but are subject to morpheme stripping, whereas words consisting of less productive or non-productive morphemes are stored in the lexicon as whole words and they don't need morpheme stripping (Gleason and Ratner, 1998, p. 164).

V. Restrictions on productivity

Viewing the vitality of the English language in terms of the influx of words created by various word formation processes, one is tempted to think that every conceivable word which can be formed is allowed, yet some factors constitute constraints on this feature. The factors that play a role in blocking or preventing a process of word formation exist in different linguistic levels, i.e., phonology, morphology, syntax, semantics and pragmatics (Katamba, 2006, p. 75; Crystal, 2003, p. 55; &Cowie, 2009, p. 17).

1.Phonological factors

Some complex words are impossible to form because they would cause difficulties for the phonetic processing, i.e., pronunciation and perception. Haspelmath (2002, p. 104) states that there is a common restriction which disallows the repetition of identical features, for instance

a- Pay….payee
b- Free…*freeee(repetition of the vowel /i:/ is not allowed).

2. Morphological factors

Katamba (2006, p. 77) mentions that the morphological properties of a base may prevent the application of certain morphological rules. English morphemes act differently from foreign morphemes. For instance, the morpheme (-ant) is attached to bases of French origin as in *defendant*, *assistant*, *merchant* etc.

Another restriction is manifested in the velar softening rule wherein /k/ changes to /s/ in words of Latin and French origins when a suffix that starts with a non-low front vowel is attached to the base, for instance /ɪ /, as in

a- Cynic …..cynicism
b- Critic …..criticism

The velar softening rule only influences words with Romance roots, whereas words that have no Romance roots don't get affected as in Isaac….Isaacism/ aɪzəkizm/ not /*aɪzəsizm/ (Katamba, 2006, p. 77).

3. Syntactic factors

According to Bauer(2001, p. 133), the most obvious constraint is that most affixes can be attached only to particular word classes, for instance, nouns, verbs and adjectives as in –ness, -en (past participle), -est respectively.Another syntactic constraint refers to the transitivity of the base, as in the affix –able which is used with transitive verbs only and not used with intransitive verbs such as *becomable(Bauer, 2001, p. 133).

4. Semantic factors

The meaning of the affix restricts the word formation rule due to the fact that some base-affix combinations make no sense (Haspelmath, 2002, p. 105), for example the prefix de- as in(*de*activate, *de*-ice, *de*frost) makes sense, because it indicates the reversal process, whereas words like **deassassinate*, **deincinerate* make no sense and are hard to interpret, because the process can't be reversed. Another restriction is about the compound words made of adjectives + past participle (v-ed) as in

a- Short- sleeved (shirt)
b- Blue- eyed (boy)

These compounds are permitted because the root to which -ed is attached is inherently possessed by the head noun it modifies (i.e., shirt, boy). In other words, the sleeve is part of the shirt and the eye is part of the body (Katamba, 2006, p. 79-80). However, **Two-carred family*is not allowed,because it is not necessary that the car is possessed by someone.

5. Pragmatic factors

The newly created words must be useful. In some languages like German which has no female nouns in –in indicating lower animalskäferin "female beetle", wurmin "female worm" these two formations are restricted, because it is not useful to make a distinction between males and females of beetles and worms (Haspelmath, 2002, p. 105).

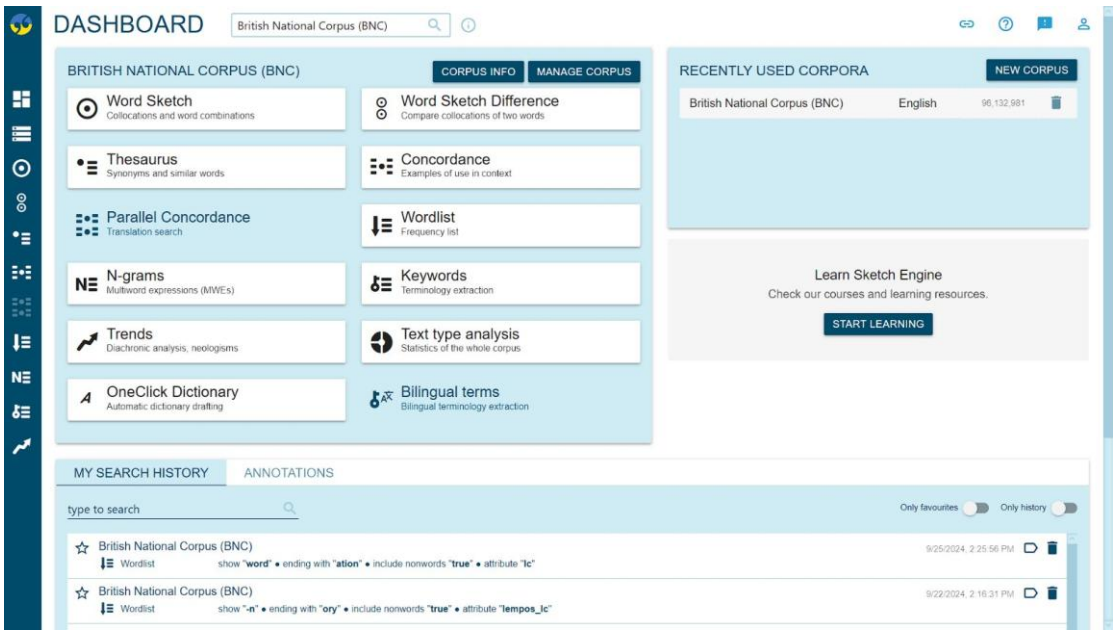
Furthermore, non-linguistic factors, i.e., aesthetic factors may inhibit some word-formation processes, as in the word "stagflation" which was coined to stand for the combination of stagnation + inflation. This word seems to have failed to stay in the language being considered as 'ugly' by some commentators (Katamba, 2006, p. 81).

VI. Data collection and description

This section addresses the rates of productivity of the suffixes used in verb-based nouns through the analysis of a large corpusin the British National Corpus (BNC)using the Sketch Engine programme (SkE)[1]. It is a web-based software system, established by Kilgarriff et al. (2004). This programme enables users to discover how language is used in huge collections of texts. Steps followed in the analysis of the data are illustrated in the following screenshots as shown below:

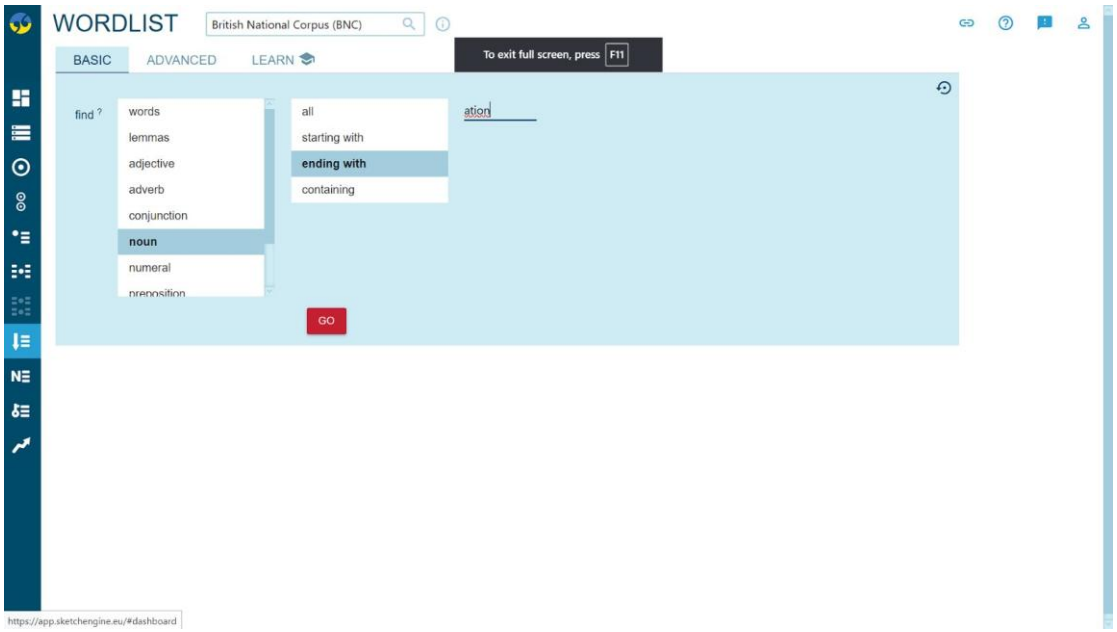
Screenshot (1) Sketch engine programme interface

[1]This corpus software is available at www.sketchengine.co.uk.



From among the tools available in the programme, the "wordlist" option has been chosen as it provides the frequencies of words and word bits, i.e., affixes, which the research paper is all about as shown in screenshot (2). A frequency list is one of the basic concepts underpinning corpus studies (Baker, 2006). It shows a word list of the occurrences of each word in the corpus under study. These frequencies have been manually cleaned.

Screenshot (2) wordlist option (word category and place of affixes)



The data which consists of (27183) words has been manually cleaned to remove the irrelevant words. The relevant instances whose type frequency has been considered are (4456) words. Bauer's (2001) model is adopted to determine the productivity of thesuffixes. Nineteen noun-forming affixes, whose meanings and origins have been verified in the Oxford English Dictionary 5thedition (2002), are investigated to check their productivity as listed in the tables(1, 2, 3, 4, and 5) below:

Table (1) Suffixes of action and states (verb+ suffix =noun)

No.	Suffix	Meaning	Origin	Example
1	-ade	Continuous action	Latin	Blockade
2	-age	Function, action, fee or charge, residence or collectivity	French	Package
3	-al	Action	Latin	Revival
4	-ive	Having a tendency towards	Latin	active
5	-ation	State, action , institution	Latin	Exploration Organization
6	-ing	Activity, result of activity	Anglo-Saxon	Driving, building
7	-ion	State or action	Latin	Action
8	-ment	Action, state	French	accomplishment
9	-sis	process of, result of	Greek	Analysis
10	-t	The act of	Anglo-Saxon	Complaint
11	-th	State, process	Anglo-Saxon	Growth, stealth
12	-ure	Act of, result of	Latin	Failure, closure
13	-y	Action, state	Latin	Recovery

Table (2) Agentive and instrumental suffixes (verb+ suffix =noun)

No.	Suffix	Meaning	Origin	Example
1	-ant	Agentive and instrumental	Latin	Inhabitant
2	-ent	Agentive and instrumental	Latin	Deterrent
3	-ist, -st	Agentive and instrumental Doer of the action	Greek	Typist, Analyst

Table (3) Passive suffix (verb + suffix = noun)

No.	Suffix	Meaning	Origin	Example
1	-ee	Passive suffix, e.g. one who is drafted	Latin	Draftee, addressee payee

Table (4) Suffixes of status and domain (verb+ suffix= noun)

No.	Suffix	Meaning	Origin	Example
1	-ery	Behavior, place of activity or abode , collectivity	Latin	Delivery

Table (5) Suffixes of location (verb+ suffix= noun)

No.	Suffix	Meaning	Origin	Example
1	-ory	A place	Latin	Observatory

(Al-Ubaidy, 2003, pp. 23-24).

VII. Data analysis

This section discusses the frequencies and percentages of the noun-forming suffixes found in the chosen corpus for the study. The analysis of the data has been tabulated as shown below:

Table (6) frequencies and percentages of noun-forming suffixes (verb+ suffix=noun)

No.	Suffix	frequency	%
1	-ade	104	2.32
2	-age	87	1.95
3	-al	205	4.60
4	-ive	101	2.26
5	-ation	1600	35.90
6	-ing	912	20.46
7	-ion	335	7.51
8	-ment	459	10.30
9	-sis	30	0.67
10	-t	6	0.13
11	-th	9	0.20
12	-ure	43	0.96
13	-y	18	0.40
14	-ant	123	2.76
15	-ent	50	1.12
16	-ist, -st	227	5.09
17	-ee	73	1.63
18	-ery	47	1.05
19	-ory	27	0.60
Total		4456	100%

Given the statistics, it has become clear that the productivity of the suffixes used in the verb-based nouns differs drastically as someshow more productivity than most others. This is manifested in the use of –ation, where (1600) cases have been found in the data representing (35.90%) of all the cases, i.e., one-third of the whole chosen corpus. That is due tothe fact that this suffix is widely used in the technical and academic terms. This widespread use is attributed to certain factors such as clarity where the suffix –ation refers to a clear grammatical function that wherever a word ends with this suffix, it is taken to be a noun. The second factor is abstraction, i.e., words ending with this suffix tend to express abstract ideas found in different contexts such as technology, politics and social sciences. The third is the formality factor, where nouns ending with the suffix –ation express a more formal tone, which prompts speakers and writers to prefer nouns ending with the suffix –ation in the discussion of abstract subjects, ideas and processes. Furthermore, the suffix –ation tends to be regular withverbs ending in –ate or –ize as in *communicate* and *generalize* to make nouns without changing the base of the word. This makes it easy to apply to a vast number of words.

Second in line comes the suffix (-ing) which occurred (912) times in the nominalization of verb-based nouns, standing for (20.46%).It represents one-fifth of all the cases. The gerund is quite commonly used functioning as a noun-likein the position of the subject, object or even the complement in the sentences, which provides flexibility in use.

As for the least productive suffixes or even non-productive ones, it has been shown that the suffixes (-t and –th) score the minimum in the whole corpus (6, 9 cases respectively) standing for (0.13%, 0.20%) due to the rarity of the words having them compared to other suffixes. The suffixes –t and –th were more commonly used in older English forms and with Germanic roots, yet they havelost productivity as English evolved and tended to preferother suffixes like (-ation, -ing, -ment) i.e., there are other competing forms used in the nominalization process.

As for the other nominalization suffixes, i.e. (-ion), (-ist,–st), (-al), they seem to have been moderately productive standing for (7.51%, 5.09%, and 4.60%) respectively. Thus, the role of productivity has become clear as it contributes to the flexibility in language use and expansion of the vocabulary enabling speakers and hearers to communicate more efficiently.

VIII. Conclusion

Productivity is a property of human language manifested in different linguistic levels. One of these levels is morphology. It helps generate new words through the addition of affixes so that language adapts to the new concepts, which facilitates communication and boosts language growth.

Given the results, it has been concluded that the productivity of suffixes used in verb-based nouns is relative to the frequencies of use of these suffixes. The most highly used suffix is – ation and the second most highly utilized one is –ingfunctioning as a gerund.

As for the non-productive suffixes, they seem to have a very low level of frequency in the whole corpus indicating that these suffixes have been fossilized. This case is represented by the suffixes –t and –th as they no longer serve as a pattern for generating new nouns, rather some other alternative suffixes took over that function. Thus, this study has shown that suffixes exhibit different levels of productivity and consequently different rates of contribution in the adaptability of the language to the new concepts and in the flexibility of use.

References

[1] Aitcheson, J., (1992). Linguistics. London: Hodder and Stoughton.
[2] Aitcheson, J., (2012). Words in the Mind: An Introduction to the Mental Lexicon (4th ed). UK: Blackwell Publishing Ltd.
[3] Al-Ubaidy, S. (2003). Morphological Transparency in English and Arabic: A contrastive study. Unpublished thesis. University of Baghdad, College of Arts. 23-24.
[4] Bauer, L. (2001). Morphological Productivity. Cambridge: Cambridge University Press.
[5] Baker, P. (2006). Using Corpora in Discourse Analysis. London: Continuum.
[6] Boij, G., (2005). The Grammar of Words. Oxford: Oxford University Press.
[7] Brown, E. K., & Miller, J. E.(1980). Syntax: A Linguistic Introduction to Sentence Structure. London: Hutchinson.
[8] Cowie, A. P. (2009). Semantics. Oxford: Oxford University Press.
[9] Crystal, D., (2003). A Dictionary of Linguistics and Phonetics (5th ed). UK: Blackwell Publishing Ltd.
[10] Gaskell, M. G. (2007). The Oxford Handbook of Psycholinguistics. Oxford: Oxford University Press.
[11] Gleason, J. B. and Ratner N. B. (1993). Psycholinguistics(2nd ed). United States: Thomson Learning
[12] Haspelmath, M. (2002). Understanding Morphology. Oxford: Oxford University Press.
[13] Katamba, F. (2006). Morphology(2nd ed). New York: Palgrave Macmillan.

[14] Kilgarriff, A., Rychly, P., Smrz, P., &Tugwell, D. (2004). The sketch engine. Information Technology, 105-116.
[15] Shorter Oxford English Dictionary on Historical Principles(5th ed) (2002). Oxford: Oxford University Press.
[16] Roach, P. (2009). English Phonetics and Phonology(4th ed). Cambridge: Cambridge University Press.
[17] Shopen, T. (1995). Language Typology and Syntactic Description: Grammatical Categories and the Lexicon. Cambridge: Cambridge University Press.
[18] Trask, R. L. (1999). Language: The Basics(2nd ed). London: Routledge.
[19] Yule, G. (2010). The Study of Language(4th ed). Cambridge: Cambridge University Press.