

# A Corpus-based Analysis of English Suffix –*esque*

Wenyu Liu

School of Foreign Languages, Dalian University of Technology, Dalian, China  
Email: wenyulaoshi@gmail.com

Haibo Shen

School of Foreign Languages, Dalian University of Technology, Dalian, China  
Email: haibo0607@yahoo.com.cn

**Abstract**—Affixation is one of the most common and productive ways to form new words in English and forming words through suffix, an affix that is placed after a base, and is a significant component of affixation. Although numerous studies on common suffixes have been conducted, relatively little attention has been paid to the seemingly less used affixes, say the suffix –*esque*. A thorough analysis is conducted of the properties of the suffix –*esque*, specifically the etymology, the semantic meaning, the formation rules, the orthographic properties, the productivity and the stylistic tendency with the help of the Corpus of Contemporary American English (COCA). The findings are helpful in English teaching or learning.

**Index Terms**—suffix, –*esque*, properties, COCA

## I. INTRODUCTION

### A. Research Background

Morphology in linguistics is always referred to as “the systematic study of morpheme” (Hu, 2001, p. 61). It studies the internal structure of words, and the rules by which words are formed (Hu, 2001, p. 61). So the study of morphology can help people better understand languages and also the nature of language. With regard to language learning, the study on morphology and its results can greatly benefit the language learners, particular in learning words. That’s why study on morphology by linguists from all over the world never ceased during the past two millennia.

One of the major research focuses in morphology is the word formation. Numerous linguists from both home and abroad have conducted investigations into the word formation rules in different languages particularly the affixation of English words. Nevertheless, former studies always focused on the common affixes, for example –*y*, –*ness*, –*ful*, –*ic/ical*, etc. Relatively little attention has been paid to the seemingly less used affixes, say the suffix –*esque*.

### B. Research Purpose

It is generally accepted that learning new words with the help of affixes is a much easier and more efficient way in English teaching and learning. Thence the author conducts a thorough investigation into the etymology, the semantic meaning, the formation rules, the orthographic properties, the productivity and the stylistic tendency of the English suffix –*esque* with the help of the corpus in order to elaborate the properties of this suffix. The findings may be helpful in teaching or learning English.

## II. LITERATURE REVIEW

### A. Affixation and Suffix

Affixation is one of the most common and productive ways to form new words in English. It is so important that almost all linguists will first define the term of affix before giving an introduction on morphology or word formation rules. As is defined by Hu Zhuanglin, affix is “a collective term for the type of morpheme that can be used only when added to another morpheme (the root or stem)” (Hu, 2001, p. 62). Hu gives another definition in his book *Linguistics: An Advanced Course Book* as that “affixes are forms that are attached to words or word elements to modify meaning or function” (Hu, 2002, p. 132). Another linguist Ingo Plag defines affix as “the cover term for all bound morphemes that attach to roots”, while roots are the central meaningful elements of words that some bound morphemes “must always be attached” to (Plag, 2002, p. 13). In his lectures, Andrew Spencer gives a definition of affix that it is “morphophonological element added to the right (suffix) or left (prefix) of a base”.

Despite the subtle variance in these definitions given by different linguists, another term base should be introduced in order to understand affix. According to Hu Zhuanglin, base is “a form to which affixes of any kind can be added” (Hu, 2002, p. 132).

Considering their syntactic and semantic properties, affixes can be classified into two types, i.e. inflectional affixes

and derivational affixes. As is pointed out by Hu Zhuanglin, “inflectional affixes are generally less productive than derivational affixes”, and inflectional affixes “often only add a minute or delicate grammatical function to the stem” (Hu, 2001, p. 63). Thus affixes in this paper refer to derivational affixes only. Meanwhile, according to their position with regard to the base, affixes can also be classified into three categories, namely, prefix, suffix, and infix. So a suffix is an affix that is placed after a base. According to the syntactical category of the derivatives, common suffixes can be further classified into nominal suffixes like *-ness*, verbal suffixes like *-ize*, adjectival suffixes like *-ive*, and adverbial suffixes like *-ly*.

### B. Productivity

Although all affixes can be used to form new words, some are more commonly used while some are not. So the term of productivity is proposed in order to explain this phenomenon. Plag defined productivity as “the property of an affix to be used to coin new complex words” (Plag, 2002, p. 44). Thus some affixes can be said to be productive since they are more frequently used to generate new words, and to the contrast some are less productive.

As pointed out above, some affixes are more productive than the other, which implies that there must be some methods to measure the productivity of affixes. In his book *Word-formation in English*, Plag offers four different measures. First, “the productivity of an affix can be discerned by counting the number of attested different words with that affix at a given point in time” which is called the type frequency of an affix (Plag, 2002, p. 52). The more derivatives are attested at that point in time, the more productive an affix is. Second, instead of counting the number of type frequency of an affix at a given point in time, the productivity of an affix can be discerned by counting “those derivatives that were newly coined in a given period, the so-called neologisms” (Plag, 2002, p. 52). Similarly, the more derivatives are attested in that period, the more productive an affix is. Thanks to the fast development of technology particular computer science, investigations into languages especially words and affixes can make use of corpora to make the work easier and the conclusions more convincing. Therefore, the other two measures involve the use of corpora. So, third, the productivity of an affix can be measured by counting the number of types (i.e. the number of different words) with a given affix and their token frequency in a proper corpus (Plag, 2002, p. 53). If there are “large numbers of low-frequency words and small numbers of high-frequency words” derived from a given affix, the affix is productive (Plag, 2002, p. 54). Fourth, the productivity of an affix can also be measured by counting “ratio of the number of hapaxes with a given affix and the number of all tokens containing that affix”. Hapaxes here simply refer to “words that occur only once in a given corpus”. “The higher the number of hapaxes with a given affix”, the more productive the affix is.

## III. METHODOLOGY

### A. Data

The paper intends to probe into the suffix *-esque* with the help of the Corpus of Contemporary American English (COCA). So the author extracted all the derivatives that are derived from the suffix *-esque* in COCA, sorted and made a statistical analysis of these words.

### B. Instrument

In order to observe the real use of the suffix *-esque*, and thus find its formation rules and measure its productivity, the Corpus of Contemporary American English (COCA) is used to gather raw materials. COCA is the largest freely-available corpus of English, and the only large and balanced corpus of American English. It was created by Mark Davies of Brigham Young University in 2008, and it is now used by tens of thousands of users every month (linguists, teachers, translators, and other researchers).

The corpus contains more than 410 million words of text and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990-2010 and the corpus is also updated once or twice a year (the most recent texts are from the summer of 2010). Owing to its design, it is perhaps the only corpus of English that is suitable for looking at current, ongoing changes in the language.

The interface allows its users to search for exact words or phrases, wildcards, lemmas, part of speech, or any combinations of these. Users can search for surrounding words (collocates) within a ten-word window (e.g. all nouns somewhere near *faint*, all adjectives near *woman*, or all verbs near *feelings*), which often provides good insight into the meaning and use of a word. (quoted from the home page of COCA)

### C. Procedure

#### 1. Obtain a wordlist of all the words end with *esque*

After entering COCA, typing “-esque” into the column of “search string” and then click the button “search”, the program offers a list of all the words end with *esque* in the right column. Altogether there are 714 types (the number of different words in the corpus) and 5375 tokens (the overall number of words in the corpus) in the wordlist.

Considering *-esque* is an adjectival suffix and has little inflectional changes, all the derivatives of *-esque* are included in the wordlist.

#### 2. Eliminate the undesired words

The words in the list are not all the derivatives of *-esque*, for the corpus can only extract words with the string of letters corresponding to the suffix including words only sharing the string of letters but not the derivatives, namely simplex words accidentally end with *esque*. Therefore, it is necessary to search each entry of the list and eliminate the undesired words, namely, words which do not belong to the derivatives of the suffix of *-esque*.

Altogether there are 25 different undesired words, including 15 proper names end with the string of letters of *esque*, 5 nouns accidentally end with *esque*, 4 words that are not derived from affixation (specifically speaking, if assume these words are derivatives of the suffix of *-esque* in the first place, the meanings of these derivatives have nothing to do with meanings of their bases; thus it is reasonable and better not to treat them as derivatives of *-esque*), and also 1 word that is not English. To make it technical, 25 types and 732 tokens are eliminated from the wordlist. Thus, the modified wordlist of derivatives from *-esque* possesses 689 types and 4643 tokens.

### 3. Observe each entry to revise the wordlist

As long as there are any variances in the forms of the words, the corpus will treat them as different entries or types in the wordlist. However, this kind of categorization sometimes may seem too rash. So with further observation and assessment, the wordlist was refined and revised.

In some cases, several words have exactly the same spelling; except for some with a hyphen between the base and the suffix, while the others without, for example, *Chaplin-esque* and *Chaplinsque*. Such words in this paper are treated as same ones, thus belong to one type.

In some cases, for the suffix *-esque* begins with the letter *e*, when the suffix is attached to a base ending with the letter *e*, the letter *e* in the base will be deleted, for example, the derivative of the base *picture* becomes *picturesque*. To the contrast, the letter *e* at the end of some base words survives when the suffix is attached to them, for example, *palace* to *palace-esque*. What's more, sometimes both forms, with the letter *e* and without, exist at the same time, for example, *google* to *googlesque* or *google-esque*. Considering the two forms, with the letter *e* or without, possess exactly the same semantic meaning; they are regarded as one type in the wordlist.

Thence, in the revised wordlist, there are 642 types and 4643 tokens. The types and tokens of the three versions of wordlist are elaborated in the following table.

TABLE 1.  
VERSIONS OF WORDLIST

version of wordlist	type	token
original	714	5375
modified	689	4643
revised	642	4643

4. Observe each entry of the wordlist to summarize the semantic meanings and the possible formation rules of the suffix *-esque*, and to analyze the orthographic properties of the derivatives of the suffix (elaborated in detail in section 4)

### 5. Calculate the productivity of the suffix *-esque*

As is mentioned in section II, there are four common measures of the productivity of affixes. In this paper, the author adopts the fourth measure to calculate the productivity of the suffix *-esque* after careful comparison of their advantages and disadvantages. The productivity is thus calculated according to the formula as following given by Plag (in which P stands for productivity,  $n_1^{\text{aff}}$  for the number of hapaxes with a given affix, and  $N^{\text{aff}}$  stands for the number of all tokens with that affix):

$$P = n_1^{\text{aff}} / N^{\text{aff}}$$

(Plag, 2002, p. 54-57)

### 6. Observe the stylistic tendency of the suffix *-esque*

COCA allows its users compare the frequency of words by genre, namely, it allows comparisons of frequency between different contexts like spoken, fiction, popular magazines, newspapers, and academic. With this program of COCA, the author compares the frequency of the suffix *-esque* between different genres of spoken, fiction, popular magazines, newspapers, and academic.

## IV. FINDINGS AND DISCUSSION

### A. Etymology

According to Wiktionary, a free on-line dictionary, the suffix *-esque* originates from the French suffix *-esque* whose semantic meaning is similar to the English suffix *-ish* or *-ic*. If dated further back, the French suffix originates from the Italian suffix *-esco* which means like, in the manner of.

### B. Semantic Meanings

According to Macmillan English Dictionary (a free English dictionary online created by teams of lexicographers who studied real spoken and written text, showing exactly how and when words are used today), there are mainly two semantic meanings of the suffix *-esque*, which was attested by the derivatives in the wordlist extracted from COCA.

The suffix *-esque* is always used with proper names, particular names of some people or places to make adjectives describing things relating to that person or their work or that place, for example, *Adamesque*, *Broadway-esque*, *Chicago-esque*, *Chomskyesque*, etc.

The suffix *-esque* is also used with some nouns to make adjectives describing things that have a particular quality relating to that noun, for example, *champagne-esque*, *hell-esque*, *king-esque*, *picturesque*, etc.

Regarding to the derivatives extracted from COCA, there are 124 types and 3059 tokens whose bases are nouns, and 501 types and 1560 tokens whose bases are proper names, as is shown in table 2.

TABLE 2.  
DISTRIBUTION OF THE BASES OF *-ESQUE*

base	type	token	Percentage of types (%)
nouns	124	3059	19.32
proper names	501	1560	78.04
adjectives	11	17	1.71
Numbers (actually nouns)	6	7	0.93
total	642	4643	100

### C. Formation Rules

According to Macmillan English Dictionary, the bases of the derivatives of the suffix of *-esque* are always nouns, either common nouns or proper names of people or places. However, as is indicated by table 2, there are other forms of the bases apart from the above mentioned nouns and proper names. After examining all entries in the revised wordlist extracted from COCA, 11 entries with adjectival bases and 6 entries with numeral bases are attested. With regard to the semantic meanings in the context in which they are uttered or written, the semantic meanings of the derivatives are obviously related to those of their adjectival bases. Thus, it is not likely that the derivatives with adjectival bases are incorrectly used. In addition, the seemingly numeral bases like *007* and *1970s* are actually treated as nouns considering their semantic meanings in the context, which suggests such cases should not be treated as abnormal.

Therefore, it is reasonable to draw the conclusion that the suffix *-esque* has loose requirements for the syntactical categories of the base, while nouns and proper names are more common (altogether accounting for 97.36% of all uses) and adjectives sometimes are also acceptable (accounting only 2.64% of all uses). What's more, it is better to note that the most common used base words of the suffix *-esque* are proper names, either names of people or places or even festivals, for this category accounts for 78.04% according to the types of the derivatives.

### D. Orthographic Properties of the Derivatives of the Suffix

#### 1. Hyphen between the Base and the Suffix

As is mentioned in section 3.3, some derivatives add a hyphen between the base and the suffix *-esque* (with 294 types, accounting for 42.67%), for example, *Buddha-esque*; while some directly add the suffix to the end of the base without hyphen (with 395 types, accounting for 57.33%), for example, *picturesque*. Furthermore, derivatives both with and without hyphen sometimes also exist, for example, *Chaplin-esque* and *Chaplinsque*. Therefore, it is sound to suggest that hyphen is not a compulsory element in word formation of the suffix *-esque*, and one can add (or not) a hyphen at the end of the base as he or she likes.

#### 2. The Letter *e* at the End of the Base

Among all the derivatives of the suffix *-esque* in COCA, there are altogether 261 derivatives whose bases end with the letter *e*, and 224 delete the letter *e*, accounting for 85.82%. The figures suggest that there is high tendency that the letter *e* at the end of a base will be deleted when the suffix *-esque* is attached to the base.

Furthermore, among the remaining 37 derivatives in which the letter *e* at the end of the bases survive, 32 have a hyphen between the base and the suffix, accounting for 86.49% (as is shown in table 3). This indicates that there is a high tendency that a hyphen will be added between the base and the suffix in the rare cases where the letter *e* at the end of the bases survive.

TABLE 3.  
LETTER *E* AT THE END OF THE BASE

	number of derivatives	percentage of all (%)	percentage of cases where <i>e</i> survives (%)
<i>e</i> deleted	224	85.82	--
<i>e</i> survives	with a hyphen	32	12.26
	without a hyphen	5	1.92
base end with <i>e</i>	261	100	

### E. Productivity

The hapax frequency of the suffix *-esque* in COCA is 455 and the token frequency is 4643, as is shown in table 4. Thus according to the formula  $P = n_1^{\text{aff}} / N^{\text{aff}}$  (in which  $P$  stands for productivity,  $n_1^{\text{aff}}$  for the number of hapaxes with a given affix, and  $N^{\text{aff}}$  stands for the number of all tokens with that affix), the productivity of the suffix *-esque* is 9.80%, which is relatively high according to Plag. This suggests that the suffix *-esque* is productive.

TABLE 4.  
FREQUENCIES

item	number
type frequency	689
hapax frequency	455
token frequency	4643
productivity	0.097997

### F. Stylistic Tendency

Clicking the button “chart”, a chart of comparison of the frequency of the suffix *-esque* between different genres of spoken, fiction, popular magazines, newspapers, and academic is obtained. The chart shows that there are 454 tokens of derivatives of *-esque* in spoken context, accounting for 8.24%; 1004 tokens in fictions, accounting for 18.22%; 1562 tokens in magazines, accounting for 28.34%; 1226 tokens in newspapers, accounting for 22.25%; 1265 in academic works, accounting for 22.95% (as is shown in table 5).

TABLE 5.  
STYLISTIC TENDENCY

section	spoken	fiction	magazine	newspaper	academic
frequency	454	1004	1562	1226	1265
Percentage (%)	8.24	18.22	28.34	22.25	22.95

Here one thing must be pointed out before any conclusions are drawn. Since COCA can only provides statistics based on the raw linguistic materials which includes the words end with the string of letters of *esque*, the frequency of derivatives of the suffix *-esque* extracted from COCA and listed in table 4 is also based on the raw materials. Considering the small number of words that are not derivatives compared with the whole picture (specifically 732 against 5375), it is reasonable to observe the stylistic tendency of the suffix *-esque* on the basis of the raw materials, and the results thus calculated are relatively approximate to the real ones. Therefore the conclusions thus drawn boast a high reliability.

Hereby the conclusion can be safely drawn that the suffix *-esque* is more common to form new words in formal contexts, for example, in fictions, magazines, newspapers and also academic works; while it is used relatively less in spoken context.

## V. CONCLUSIONS AND SUGGESTIONS

### A. Conclusions

On the basis of linguistic materials extracted from COCA and through careful statistic analysis, the etymology, semantic meanings, formation rules, orthographic properties, productivity and stylistic tendency of *-esque* are analyzed in detail and the following conclusions are drawn:

First, the suffix *-esque* originates from the French suffix *-esque* which means like, in the manner of.

Second, the suffix *-esque* mainly possesses two semantic meanings, one is used with proper names, particular names of some people or places to make adjectives describing things relating to that person or their work or that place, and the other is used with some nouns to make adjectives describing things that have a particular quality relating to that noun.

Third, the suffix *-esque* has loose requirements for the syntactical categories of the base, while nouns and proper names are more common and adjectives sometimes are also acceptable. What's more, the most common used base words of the suffix *-esque* are proper names, either names of people or places or even festivals.

Fourth, hyphen is not a compulsory element in word formation of the suffix *-esque*, and the letter *e* at the end of a base tends to be deleted when the suffix *-esque* is attached to the base. Meanwhile, there is a high tendency that a hyphen will be added between the base and the suffix in the rare cases where the letter *e* at the end of the bases survive.

Fifth, the suffix *-esque* is productive.

Sixth, the suffix *-esque* is more common to form new words in formal contexts, for example, in fictions, magazines, newspapers and also academic works; while it is used relatively less in spoken context.

### B. Suggestions for Further Study

The analysis and research of the suffix *-esque* in this paper are based on a single corpus, COCA, which is of American English only. Thus the conclusions drawn on the basis of linguistic materials extracted from COCA in this paper may be constrained and confined. Therefore, the author offers a suggestion that further studies may be conducted based on British National Corpus (BNC) or on the two corpora.

## REFERENCES

- [1] Corpus of Contemporary American English, URL: <http://corpus.byu.edu/coca/> (accessed 20/12/2010).
- [2] Plag, I. (2002). *Word-formation in English*. Great Britain: Cambridge University Press.
- [3] Bauer, L. (1988). *Introduction Linguistic Morphology*. Great Britain: Edinburgh University Press.

- [4] Macmillan English Dictionary, URL: <http://www.macmillandictionary.com/dictionary/american/esque> (accessed 15/12/2010).
- [5] Wiktionary, URL: <http://en.wiktionary.org/wiki/-esque> (accessed 15/12/2010).
- [6] Hu, Z. L. (2001). *Linguistics: A Course Book*. Beijing: Peking University Press.
- [7] Hu, Z. L. (2002). *Linguistics: An Advanced Course Book*. Beijing: Peking University Press.

**Wenyu Liu** received his Ph.D. from Dalian University of Technology and now is a Professor in the School of Foreign Languages at Dalian University of Technology. He has been teaching since 1993. From 2009 to 2010, he worked at the University of Minnesota Twin cities, the U.S.A. as a Fulbright research scholar. His research interests are teaching English as a second or foreign language, computer assisted language learning and neurolinguistics.

**Haibo Shen** is now studying in Dalian University of Technology as a master candidate. Her research interests are ESL learning strategy, computer assisted language learning and neurolinguistics.