Tampere University

Emilia Vainionpää

THE USE OF *-ESQUE* AS AN ADJECTIVAL SUFFIX WITH PROPER NOUNS IN THE IWEB CORPUS

Faculty of Information Technology and Communication Sciences
Bachelor's Thesis
April 2023

TIIVISTELMÄ

Emilia Vainionpää: Adjektiivijohdin *-esque:*n käyttö erisnimien kanssa iWeb-korpuksessa
Kandidaatintutkielma
Tampereen yliopisto
Kielten kandidaattiohjelma
Huhtikuu 2023

Tässä kandidaatintutkielmassa tarkastelen miten adjektiivijohdinta *-esque* on käytetty yhdessä erisnimien kanssa englanninkielisessä iWeb corpus -aineistossa (I. Intelligent Web-based Corpus). Adjektiivipääte *-esque* tekee substantiiveista adjektiiveja, joiden merkitys on: ”jonkin tyylillä tai imitoiden x:ää”. Yleisin erisnimestä *-esque:*lla johdettu adjektiivi iWeb-korpuksessa on *Kafkaesque*, joka on vakiintunut viittaamaan Franz Kafkan tuotannossa esiintyvän todellisuuden vääristyneisyyteen ja painajaismaisuuteen.

Tutkimuksen aineistoksi valikoitui iWeb-korpus, koska halusin tutkia hiljattain käytettyjä adjektiiveja. Vuonna 2017 luotu IWeb-korpus koostuu yli 14 miljardista sanasta yli 94 tuhannelta nettisivustolta, ja on yksi suurimmista englanninkielisistä korpuksista. Kirjallisuuskatsauksessa esittelen morfologiaa, sananmuodostusta, uudissanoja, *-esque*-päätettä ja affiksien produktiivisuutta sekä kerron lyhyesti korpuslingvistiikasta. Metodiosiossa kerron käyttämistäni korpushauista ja iWeb-korpuksen erilaisista hakumahdollisuuksista.

Tarkastelun kohteena oli etenkin ”*esque”-merkkijonohaun sadan yleisimmän sanan lista, mutta tein myös tarkentavia hakuja selvittämään esimerkiksi, keihin samaa sukunimeä kantaviin henkilöihin *-esque*-päätteisellä adjektiivilla viitattiin. Tutkin myös, onko adjektiiveissa yleisempää käyttää väliviivaa vai jättää se pois, ja esittelen korpusmateriaalista löytyviä erikoisia sanamuotoja. Analyysiosiossa käsittelen sukunimistä johdettujen adjektiivien lisäksi myös taiteelliseen mediaan (kuten kirjoihin, peleihin tai fiktiivisiin hahmoihin) liittyvistä erisnimistä johdettuja adjektiiveja, sekä muutamaa yhtiön nimestä muodostettua adjektiivia.

Tutkielmani johtopäätöksiä ovat, että internet-tekstejä sisältävässä iWeb-korpuksessa *-esque*-adjektiivipäätteellä on muodostettu erisnimistä tuhansia erilaisia adjektiiveja, joista muutama yleisin löytyy sanakirjasta, mutta suurin osa on uusia, tai ei koskaan sanakirjoihin päätyneitä. Sadan yleisimmän *-esque*-päätteisen sanan joukosta löytyy 26 sukunimestä luotua adjektiivia, joista suurin osa viittaa taiteilijoihin. Poliitikkojen nimiin viitataan neljässä sadasta tuloksesta, ja urheilijoiden nimet ovat harvinaisia jopa tuhannen yleisimmän tuloksen listassa. Toiseksi yleisin erisnimiryhmä on taiteelliseen mediaan liittyvät erisnimet, joita löytyy sadan joukosta kuusitoista erilaista. Jotkut näistä kuitenkin paljastuivat eri yhdyssanoista koostuviksi tulosjoukoiksi. Yhtiöiden ja niiden tuotteiden nimistä johdettuja adjektiiveja oli kolme sadasta.

Tutkimusta voisi jatkaa tutkimalla iWeb-korpuksesta esimerkiksi sitä, minkä sanojen kanssa tietyt *-esque*-päätteiset adjektiivit esiintyvät, tai tutkia adjektiivipäätettä erilaisissa internetkorpuksissa. Tutkimalla eri vuosikymmenillä internetissä suosituimpia *-esque*:sta johdettuja adjektiiveja voitaisiin selvittää, vaikuttavatko henkilöiden tai median tuotteiden tunnettuus tai kulttuurinen merkittävyys tiettynä ajanjaksona niiden yleisyyteen.

Avainsanat: esque, adjektiivipääte, uudissana, sananmuodostus, korpustutkimus, iWeb corpus

Table of Contents

1. Introduction

In the current thesis I will investigate the adjective suffix *-esque* in contemporary use via the iWeb (The Intelligent Web-based) Corpus from 2017. *The Oxford English Dicti*onary defines the suffix as 'resembling the style partaking of the characteristics of' the base word (OED, s.v. *-esque,* suffix). I will exclusively examine adjectives that are formed of a proper name and the suffix, and not established adjectives that consist of a common noun and end in the *-esque* suffix, such as *picturesque* or *grotesque.* The best known of the *-esque* adjectives derived from proper names is *Kafkaesque* which is defined in *Oxford Living Dictionaries* as 'characteristic or reminiscent of the oppressive or nightmarish qualities of Franz Kafka's fictional world', but I want to examine what other proper names the suffix has recently been affixed to in internet texts. In using the iWeb corpus, I will not only observe the language use of language experts, such as linguists or journalists, but also how ordinary language users have created new adjectives. The prediction is that in an internet corpus such as this, with not only text from news sites, but varied types of websites, the used *-esque* adjectives are numerous and creative like the writers who use or conceive of them.

I aim to demonstrate and describe what kind of proper names the suffix *-esque* is added to using the corpus and show how the iWeb corpus can be utilised to examine the suffix. I will discuss how the adjectival suffix *-esque* has been used to form adjectives with proper names in the iWeb corpus data and explain what can be observed from the way the proper-name-derived adjectives are formed. I will additionally consider whether there is anything to be said about the use of the hyphen in the adjectives, and whether there seems to be any pattern to using the hyphen (e.g. *Beatles-esque*), not using it (e.g. *Beatlesesque*), or blending the base and the suffix (e.g. *Beatlesque*).

The most recent study on the *-esque* suffix was made by W. Liu and H. Shen using the COCA (Corpus of Contemporary American English) in 2012, but they focused generally more on the statistics rather than the various nouns used with the suffix. Furthermore, the COCA consists of one billion words as of 2023, but it is likely that in a far more extensive corpus the findings will be

different. There are no studies conducted on the suffix *-esque* specifically in language used on the internet so far, nor using a corpus as vast as the 14-billion-word iWeb corpus. There are many past studies and books that briefly mention the suffix and how it is used, but the scope of this thesis is exclusively limited to the current use in the iWeb corpus data compiled in 2017.

In the next section I will discuss the linguistic phenomena and background material that pertains to the subject at hand. In the third section, the materials and methods will be discussed, with a dedicated sub-section to the iWeb corpus. The analysis and results will comprise the fourth section, and lastly, I will conclude the thesis with a summary and suggestions for further research. Additionally, two screenshots of the corpus data can be found after the bibliography.

2. Literature review

In this section the relevant literature will be reviewed in brief. First, I will discuss morphology and word-formation, neologisms, and the *–esque* suffix itself. Secondly, I will briefly discuss the concept of corpus linguistics.

2.1 Word-formation and morphology

Morphology is the linguistic field that studies words, their structure, and how they are formed. Word-formation is a branch of morphology that deals with the creation of new words and aims at describing different ways that words are formed, for example derivation and compounding. Word-formation is often contrasted with inflection, which is the morphological process that "encode[s] grammatical categories such as tense, case, plurality, or person" (Plag 2003, 14). Inflectional affixes in English are always suffixes, that is, they only attach to the end of the base and never change the word class of the lexeme whereas derivational suffixes can, but do not always (Plag 2003, 15). To elaborate, lexeme is the term used to refer to "realizations of an abstract morphological entity" (Plag

2003, 9), whose different "grammatically specified forms . . . are called word-forms or grammatical words" (Plag et al. 2009, 90). Lexemes are usually represented with small capitals.

In morphology, the smallest part of a word that has meaning is called a morpheme, and all words consist of one or more morphemes (Plag 2003, 10). Words that cannot be divided into smaller units are called monomorphemic words. To illustrate, the words *word* and *unit* are monomorphemic words and they can be contrasted with polymorphemic, or more commonly, complex words. As an example, the word *formation* is a complex word which can be taken apart into the morphemes *form*, which is a Latinate noun meaning 'shape', and *-ation*, which is a derivational suffix that forms nouns. Morphemes can be divided into two categories: free morphemes that can occur by themselves, and bound morphemes, that must attach to a 'base' or 'root'. Plag (2003, 11) defines *base* as a 'word which an affix is attached to-, which can consist of multiple morphemes, whereas a *root* is 'a type of base that only consists of one morpheme and therefore is not further divisible'.

The morphological process of attaching affixes to bases is called affixation. Affixes are usually bound morphemes that attach to bases, but there are a few affixes that are homographic, and not immediately easy to define as bound or free morphemes (Plag 2003, 73). The three types of affixes are named according to the position they attach to in a base. Prefixes, such as *post-*, *in-* and *dis-* attach to the front of bases, infixes can be inserted in the middle of bases, and suffixes, such as *-ness*, *-ity*, and *–esque* are attached to the end of a base.

2.1.1 The suffix *–esque*: evaluation and meaning

According to the OED, the suffix *–esque* comes from French, via Italian *–esco* from medieval Latin *–iscus*, and it is suggested that in Italian, writers use *-esco* primarily with artist's names, and that "French and English writers on art have imitated this practice" (OED, s.v. *–esque*). It is a bound

morpheme, which attaches to noun bases to create adjectives. Therefore, *-esque* is a category-changing derivational suffix because it is attached to nouns to form adjectives (Plag 2003, 15).

Plag (2003, 95) reports that the suffix is attached to "common and proper nouns to convey the notion of 'in the manner or style of X'*: Chaplinesque*, *Hemingwayesque*, *picturesque, Kafkaesque*. There is a strong preference for polysyllabic base words." Additionally, according to Marchand (1969, 286), the *-esque* suffix "forms adjectives with the meaning 'having the (artistic, bizarre, picturesque) style of –'". Marchand (ibid.) also maintains that "the word **picturesque** . . . characterizes the whole semantic field of the suffix" (word bolded in original). Moreover, Moon (2008, 140-1) notes: "The suffix *-esque* . . . [occurs] after nouns or names in well-established adjectives such as *picturesque* and also *statuesque, rubenesque, junoesque,* referring to (mainly female) body types", and suggests that the suffix "is used productively after the names of writers, artists, entertainers, politicans [sic] and so on, to form adjectives which indicate or allude to qualities typical of the person mentioned" (ibid.).

The words formed of proper nouns and the *-esque* suffix have can be seen as having compositional meaning, "a function of the meaning of its parts" (Plag 2003, 21). This can be observed in dictionary entries of different *–esque*-final words, e.g. *Americanesque* is simply defined as 'of a style or manner characteristic of the United States or its inhabitants' (OED, s.v. *Americanesque*). However, there are exceptions already recorded in dictionaries, for example, *Kafkaesque*, which has acquired an idiomatic meaning of 'having a nightmarishly complex, bizarre, or illogical quality' (MW, s.v. *Kafkaesque*) and *Beatlesque*, which is defined as 'typical or redolent of the catchy, upbeat melodies, close vocal harmonies, or tight, guitar-led arrangements of early Beatles records' (OED, s.v. spec. sense *Beatlesque*).

The OED entry suggests that the words formed with the suffix "are chiefly nonce-words of a jocular character, as *cigaresque*" (OED, s.v. *–esque*). Additionally, Moon (2008, 141) speculates that in the Bank of English corpus data the more rare *-esque,* or "ad hoc" words have a somewhat

negative connotation, "and imply 'less good than' or 'imitative'", whereas the more frequently used "formations mainly have more specific meanings and often neutral or positive evaluations". Her example BoE concordance lines include such rare formations as *the appalling Kim Wilde-esque ballad,* and *narrative involving a group of Waco-esque religious oddballs*, where the negative evaluation is apparent from the use of the negatively charged words *appalling* and *oddballs* (Moon 2008, 141). Interestingly, the suffix *–ish* which has a similar use, specifically 'of or belonging to a person or thing, of the nature or character of' (OED, s.v. *–ish*, sense 2.) is recorded in the OED as having "[i]n recent colloquial and journalistic use, … become the favourite ending for forming adjectives for the nonce (esp. of a slighting or depreciatory nature) on proper names of persons, places, or things, and even on phrases" (Moon 2008, 141). These two evaluations would suggest that the suffixes *–ish* and *–esque* have had similar uses fairly recently. However, it might be noteworthy that the OED descriptions of either suffix have not been updated since 1989, although it is possible that the descriptions would not need alterations. Furthermore, evaluating connotations the *-esque* adjectives have in the short context provided in the iWeb corpus might be impossible, and indeed reading the longer context is beyond the scope of this thesis.

2.1.2 Neologisms

A *neologism* is defined in the OED as 'a word or phrase which is new to the language; one which is newly coined' (s.v. sense 1.a). Why do language users create new words? According to Plag (2003, 59-60), there are three main reasons for word-formation. The most obvious reason is to name new concepts that have not existed before, called labelling or referential function. Additionally, syntactic recategorisation is used to condense information, for example by substituting a long clause with a single word, or for stylistic reasons. Words like *sweetie*, according to Plag (2003, 60) are formed to express the user's attitude towards a person or thing. Mattiello (2017, 25) argues that lexicographers and linguists define neologisms as permanent additions to the language whereas

nonce-words or *occasionalisms* are meant for specific, one-time usage. “A new word is truly new only at the time when it enters the lexicon of a language”, and therefore after a word has been recorded in a dictionary and been in use for decades, it is perhaps no longer considered a neologism, claims Mattiello (2017, 26). Many of the *-esque* formations in the present thesis’ data can be seen as true nonce-words, not attested in any dictionaries, and perhaps never mentioned anywhere but the website from which they were retrieved into the corpus.

2.1.3 Productivity of affixes

The productivity of an affix is the property that it can be used to create new words. Some affixes can be used to produce new words, whereas others cannot. For example, the category-changing *–th* suffix is an unproductive one, because it can only be used with some adjectives to create a noun, to illustrate, adj. *strong* + suffix *-th* -> n. *strength*. In theory, the suffix *-esque* could be added to any proper noun to make a grammatical adjective, even if its meaning would be very bound to the context and its use only for one time as a nonce-word. It can be argued that some affixes are more productive than others, and that their rate of use varies over time.

Productivity is difficult to measure, and Plag (2003, 51-52) proposes some of the methods used are problematic. This is especially because the amount of words in a dictionary with a given affix does not show whether or not an affix was productive in the past or whether it is productive in current use. The productivity can more accurately be measured by counting the new words coined in a specified time period, known as neologisms. Plag (2003, 52) suggests the use of the OED in CD-form to find neologisms, since it can be searched for specific time periods and it aims to be a complete account of English vocabulary. However, Plag adds that this is not the perfect tool for the task, because it may be that some neologisms have been disregarded or not added by the dictionary editors. Definitely, even now that the dictionary is online and is being edited constantly, it is far from a complete account of the English language in use.

2.2 Corpus linguistics

"A corpus is a collection of spoken of written texts to be used for linguistic analysis and based on a specific set of design criteria influenced by its purpose and scope." Here Weisser (2015, 23) uses a simple definition of the word, which includes any text collection compiled "systematically" for the investigation of linguistic phenomena. This definition therefore includes even small collections of texts, as well as vast ones consisting of millions of texts. Weisser (ibid.) further clarifies that most corpora in current use are computerised. Corpora exist in several different genres, by what types of texts they consist of, like internet news (e.g. News on the Web corpus (NOW)), British Parliament texts in the Hansard Corpus, or historical corpora with various text types divided by decade (like Early English Books Online (EEBO)).

Corpus linguistics is a data analysis methodology, and in fact, Weisser (martinweisser.org) critises the use of *corpus linguistics* as a term, and prefers to use the term *corpus-based linguistics* to clarify that it is a methodology to be used in any branch of linguistics, and not a branch of linguistics in itself. He adds that the focus when using corpus-based linguistics should be on the language studied, not the corpus itself. Corpora can be utilised in various types of linguistic research, or even just to determine how a word is used for the sake of writing a better paper in a foreign language. For the latter purpose, a collocation list, which shows what words collocate, which means they are frequently found together, can be very useful.

Corpora can be divided into two categories by the nature of their closedness or openness. Most corpora are static, synchronic, or snapshot corpora, which exhibit the language used at a particular time, and are never added to (Weisser 2015, 37), or they can be monitor corpora like COCA and the Bank of English, which are designed to be expanded and reflect the change in language use over time. A corpus can be used to search for individual words, phrases or *lemmas* "an abstract lexical category [that consists] of a lexeme base and its inflected forms" (Weisser), and has different ways of showing the search terms for different types of research.

3. Materials and methods

This section will be dedicated to the materials used in the study, namely the corpus and dictionaries, and the methods. The basics and creation of the iWeb corpus are discussed in brief in section 3.1.1, and in section 3.2, the possibilities of the different corpus search queries and iWeb's limitations are considered.

3.1 Materials

The materials used in this thesis consist of the iWeb Corpus, available from the English Corpora website (english-corpora.org/iweb/), and searches made from it with different search terms. Dictionaries, such as the *Oxford English Dictionary*, *Oxford Living Dictionaries* and *Merriam Webster* were used to find definitions for some words for illustrative purposes.

3.1.1 The iWeb Corpus

The iWeb Corpus is one of the corpora in the English corpora website (english-corpora.org) online, and one of the three largest corpora (according to english-corpora.org), with its word count of 14 billion in over 22 million webpages, from over 94 thousand websites. The data was compiled by downloading data from Amazon's Alexa, of the top one million websites by number of users in 2017. Then it was narrowed down to include websites where most of the users are from "the US, Canada, Ireland, the UK, Australia, and New Zealand" (iWeb, english-corpora.org), to acquire text mostly in English. The data were then also narrowed down to include only websites with the word *of* (because the creators deemed most sites would have the word *of)* in them from *Google* and *Bing* searches and downloaded with custom software made by Spencer Davies in the *Go* programming language. The data was further reduced from some 200 000 sites, to its final 94 thousand, to only

include sites with enough text in them and exclude, for example, pages blocked for various reasons by the university proxy server.

The corpus is fast in comparison to many older, smaller corpora. It can be searched for any topic, for example *video games*, and then divided into smaller, more specified corpora in that manner. The iWeb corpus is argued (by its creators) to be the best corpus for language learners, especially because of its top 60,000 most frequent words list. The top list can be used in learning English, by listening to audio, watching videos, and looking at images of any of the top 60,000 words. The corpus can be searched for topics, collocating words, word clusters, concordances, and by websites.

The iWeb corpus is a synchronic corpus, although not all of the data found in it was written in the year it was compiled, or even originally written to be published on the internet because it is cited from older literature. The internet itself has been in commercial use for three decades, so some of the data may be from decades back. Therefore, it might be slightly problematic to claim that the present thesis only aims to illustrate how the *-esque* suffix was used a few years ago on the websites specifically compiled from the internet for the corpus.

3.2 Methods

My study will be mostly qualitative, but also include some statistics about the frequency of the adjectives found in the corpus searches. The data is gathered mainly via “List” corpus searches, where a search string is used to discover a list of the most frequent results for the word used. The “List” function allows for searches with thousands of hits, but in the present thesis, the most used list was the 100 and 300 most frequent hits found. The “Frequency” tab allows looking through every single concordance line for the chosen word in the “Context” tab. The key words, however, do not vertically line up in the iWeb corpus, which somewhat impedes browsing through the data.

The material in this thesis was compiled by searches from the iWeb corpus, primarily with the search term "*esque". With this search string, the search interface retrieved all possible instances where the suffix is either preceded by a space, a hyphen (or another type of punctuation mark) or where it is attached to a word or letter. Other search terms were also utilised, for example to determine whether a word is used more with or without a hyphen, by searching for specific words such as *Disneyesque* or the hyphenated alternative *Disney-esque*. By searching with the string "*-esque", it would be possible to gauge what formations are most frequent with the hyphen attached, but comparing that list to the *-esque* list would not have been effective, because it is impossible to search for a list with exactly the proper noun + *-esque* formations. Searching for the possibly miswritten or misrepresented (e.g. different in the corpus than the original website) forms, with a space between the preceding word and the suffix was possible with the string " *esque"*, but that query shows the same results as clicking the free-standing *esque* in the "**esque"* results, discussed in section 4.1.

The corpus can also be searched with a search string where the part of speech, proper noun, tag is added, but the searches "NAME esque" or "NAME *esque" do not include all of the "proper noun + esque" words. It seems that not all of the proper nouns in the corpus are tagged as proper nouns, because the search string "nn_ *-esque" for example, yields results where the noun that is tagged as a common noun (not proper noun), is inarguably a proper noun, such as *Lady Gaga-esque* referring to the musician, or *Blade Runner-esque*, referring to a film name or franchise. The corpus might not tag every noun with two or more words in it as a proper noun even if it should.

Some searches for collocating words are also done to observe whether the examined adjective is a compound proper noun, like *James Bond-esque* when the 'List' search result does not show the first part *James*. The 'Collocates' search finds how often other words are nearby the search string, but in the present thesis it was only utilised to find compounds that the 'List' cannot show. By choosing to only show the collocating words exactly to the left, it was possible to determine how

many *Michael Jackson-esque* hits there were compared to *Janet Jackson-esque* for example, without having to manually count the names from the 'Context' page in a 'List' search.

The corpus may also include problematic searches, and it is not always be possible to ascertain whether the search result is as it was created on the particular website or whether it is altered in the corpus. One such case in the search seems to be *LTvesque* wherein the corpus has substituted the letter "é" or "è" for "T". The search "* esque" has the most frequent hit as " " esque " where the quotation mark can refer to quotation marks in the original web text, but it also serves as a placeholder for other punctuation marks, like the en dash (–).

In searching a corpus, a methodological problem that can arise is that the data will always have to be processed in order to be able to analyse it effectively. In the case of the present thesis, the corpus search results were analysed by examining the words in a broader context so that words that are not formed of a proper noun and the *–esque* suffix, will not be included. Architectural styles' or artistic movements' names formed with a common noun (e.g. *plateresque*) or proper noun (*e.g. Cosmatesque,* from Roman *Cosmati* family) were likewise disregarded, because they are established styles and can be accurately described, and they have *Wikipedia* pages, unlike even the most frequent proper name + *esque* adjective, *Kafkaesque*. Some of these types of *esque* formations, like the 82nd most frequent hit, *Churrigueresque* 'relating to Jose Churriguera's (1650-1732) style of architecture or its imitations', can indeed also be used as a noun.

4. Analysis and results

Analysing and evaluating examples of actual language use from the internet is complicated, because there are no set rules on how to do it. There is an overabundance of possibilities for what to examine in the iWeb corpus with the query "**esque"*, because it provided a total of 177,158 hits, out of which 13,375 were unique. The true uniqueness of each hit as a word is questionable because some of them are miswritten forms of certain words. This can be seen for example by browsing the

bottom of the **esque* 3000 most frequent list, where there are two examples of *Frankeinsteinesque*, which is almost definitely a typographical error with the extraneous *i* in the middle. The language users in the corpus data, in this case, may be professional writers, like journalists, but they can also be writers of forum posts, comments and the like. Bloggers and internet commenters are often ordinary language users, not professionals, and there are typographical as well as grammatical errors in the data. It will most likely not be possible to distinguish what sort of an error, if any, a possible oddity in the data will be, nor whether it be a deficiency in the corpus or the original written text.

4.1 The most frequent formations and problematic cases

In the corpus search for **esque* the most frequent word ending in *–esque* was *picturesque* (80,932 hits), followed by *grotesque* (18,054 hits) and *burlesque* (12,110 hits). These adjectives and nouns are very frequent and well-known, and they will not be discussed because none of the bases are proper nouns. The fourth result, *Romanesque,* however, is formed of the adjective (or proper noun) *roman*, usually denoting an architectural style reminiscent of Ancient Rome, however, it also falls beyond the scope of this thesis as discussed in section 3.2. The fifth most frequent hit *Levesque* and the eighth *Presque* (also a French word for 'nearly/almost') were themselves proper nouns and not even adjectives. The 14th most frequent hit, *resque* must also be disregarded, because it is the name of a library for the Ruby programming language, and resembles the word *rescue*, therefore it is also a possible miswritten form of the verb or noun.

The first dilemma emerges at number (9); simply, *esque,* with 1,205 hits. Upon first inspection the keyword *esque* by itself provides a list of compound words that use the suffix as a stand-alone word, as if it were a free morpheme. There are examples such as (4) *the old 70's Johnny Rodgers esque* (an American ex-footballer), (15) *Shirley Temple esque* and (16) *classic pop Madonna esque*, where the possible reason for the lonely *esque* could be that it would simply look

strange to the writer to attach it to the name, or that the author was unsure of how to write the newly coined word. Other strange-looking examples are (19) *the "Jaws" esque* and (7) *Warcraft 3 esque*, the former about a shark thriller film and the latter a video game from 2002.

Perhaps the most interesting one of the simple *esque* results is (22) *Linda Blair- esque*, where a hyphen is used but it is attached to the surname *Blair* instead of the suffix. This is reminiscent of the Finnish way of using the hyphen to signify a compound word whose first part is a *sanaliitto (eng.* 'set phrase'), such as *Taru Sormusten Herrasta* –kirja, (*eng. a Lord of the Rings* book). However, *Linda Blair- esque* has the hyphen in the wrong position for it to be a well-formed English compound word, or even a Finnish one. Number (13) on the list is *–esque* by itself, with 527 hits, which is less than half as frequent as the suffix by itself without the hyphen. However, there seems to be an error in the corpus for this search, because it interprets "-" in the frequency results as a minus symbol, and therefore attempting to view "-*esque*" results in the corpus searching for everything that does not have the string of characters "*esque*" in it. This renders the "-*esque*" results unanalysable.

It is possible that the further down the frequency list the adjectives (and nouns) are analysed, and the less there are hits for particular words, the more difficult it will become to glean who the words refer to because the words are rarer. A writer may have coined an adjective of a person they know, or a very minor celebrity, and the context may not communicate who the referent is. Recognising proper nouns from the list is not certain, and even by inspecting the words in context, there is a possibility that some of the proper noun formations are mistakenly disregarded, either by accident or a fault in the corpus, if the noun is not correctly capitalized. In this type of analysis, the aim is to be descriptive rather than systematic and precise.

4.2 The most frequent proper noun formations

The tenth most frequent hit, the first one relevant to this thesis is *Kafkaesque* with 1032 hits, which shows that according to this 2017 internet corpus data, it is the most common proper noun + *-esque* derivation as suspected. The second most frequent is *Disney-esque* with 222 hits, and third *Bond-esque* with 212 hits. Examining further, *Disney-esque* seems to have only four tokens with Walt Disney's first name. This could be due to most of the tokens describing his work, or the work of his company, *The Walt Disney Company*, usually shortened to simply *Disney*. *Bond-esque* on the other hand is preceded by *James* in 112 of the tokens, which is more than half of them, even though the name *Bond* is solely used with the famous fictional spy.

Pythonesque is the third most frequent, but its frequency of 193 can be distorted by its three different apparent meanings. It seems to be most often used to denote something *Monty Pythonesque*, as in something in the style of the British comedy group, and written with a capital P. The second meaning for *Pythonesque* refers to the programming language, whose name is styled in the logo as *python* with lower-case letters, but it is also written with a capital P. The third meaning, however, written with a lower-case p, is not at all an adjective derived from a proper name but one denoting something snake-like.

4.2.1 Last name formations

The most accessible classification when analysing the top 100 most frequent **esque* adjectives is finding all hits where the word is formed of a real person's last name. Most of the ones in the table below are recognisable by the last name, or at least effortless to identify the referents to online.

No.	Pos in 100	Word	Freq.	Vowel	Cons.	Field	Additional Notes
1	10th	Kafkaesque	1032	V		ART	
2	16th	Disney-esque	222	V -		ART	
3	19th	Kafka-esque	190				Both versions = 1222
4	20th	Tolkienesque	176		C	ART	
5	21st	Rubenesque	158		C	ART	
6	23rd	Disneyesque	148				Both versions = 370
7	30th	Dylanesque	130		C	ART	
8	31st	Tolkien-esque	126				Both versions = 302
9	33rd	Bowie-esque	107	V -		ART	
10	34th	Reaganesque	107		C	POL	
11	35th	Tarantino-esque	107	V -		ART	
12	36th	Dali-esque	105	V -		ART	
13	43rd	Burton-esque	89		C -	ART	
14	45th	Chaplinesque	87		C	ART	
15	46th	Dylan-esque	87				Both versions = 217
16	48th	Trump-esque	86		C -	POL	
17	52nd	Escher-esque	82		C -	ART	
18	54th	Clintonesque	79		C	POL	
19	63rd	Whedonesque	67		C	ART	
20	66th	Stalinesque	64		C	POL	
21	67th	Warhol-esque	64		C -	ART	
22	76th	Dantesque	58	V		ART	Blend
23	78th	Goldberg-esque	57		C -	ART	
24	81st	Whitmanesque	56		C	ART	
25	85th	Fellini-esque	54	V -		ART	
26	86th	Picasso-esque	54	V -		ART	
27	88th	Burtonesque	53				Both versions = 142
28	93rd	Kennedy-esque	52	V -		POL	
29	95th	Pinteresque	52		C	ART	
30	98th	Capra-esque	49	V -		ART	
31	100th	Jackson-esque	49		C	ART	Michael in 25
Most frequent without hyphen/with hyphen				2/8	11/5		

Table 1. iWeb Corpus "*esque" frequency results of last name + *esque*/*-esque* formations

As can be seen in this table, there seems to be a slight preference to using the hyphen before the suffix, and a strong preference to use the hyphen when the last name ends in a vowel. There are only three occurrences in these 31 adjectives where a last name ends in a vowel, and no hyphen is added before the *-esque* suffix, the most common one, *Kafkaesque*, sixth most common, *Disneyesque,* and somewhat less common, *Dantesque,* which is a blended form, omitting the other *e* in the middle. In addition to the previously discussed first two, *Dantesque* is also attested in the OED, recorded in 1813 so it is an older word than many other adjectives in the corpus data. It is possible that adding the hyphen in most of the adjectives better clarifies what person is being referred to, or that the writers feel that two vowels from the name root and the suffix adjoining would be awkward.

Of the hundred most frequent hits with the search string **esque* 31 are words created with real people's last names. Five of them are a different iteration of the same word, appearing with a hyphen and without it. These are adjectives formed of the names Kafka, Disney, Tolkien, Burton, and Dylan. It is possible to categorise these adjectives of real people's names ending with *-esque* or *esque* in different ways by finding their referent. 21 of the hits refer to individuals that are known for their art, for example, painters, songwriters, actors, or writers. Only four hits referred to people that were presidents or political leaders: *Reaganesque*, *Clintonesque*, *Stalinesque* and *Kennedy-esque*.

Most of the referents for a word are individuals with distinct or unique names, for example, *Tarantino-esque*, *Dali-esque*, or *Warhol-esque*. In the case of common last names, there are hits where the last name has different bearers, distinguishable by the presence of the first name in the context. The 100th most frequent hit, *Jackson-esque* is a great example of this. The most frequent first name preceding *Jackson-esque* is *Michael,* referring to the singer, but there are two examples of *Janet Jackson-esque*, referring to Michael's singer sister. Other than these Jacksons, there is one example of *Wanda, Shirley, Fred, Greg, Joe,* and *Peter* each. *Wanda* is a singer and songwriter, *Shirley* was a writer, *Fred* is a former American footballer, *Greg* a mixed martial arts (MMA)

trainer, *Joe* the father of the Jackson family mentioned before, a talent manager, and *Peter* a film director, screenwriter, and producer.

Among these most frequent last name + *esque* formations, there are only references to famous people who are artists of differing fields and a few politicians, but there are no other types of celebrities. The first and most common adjective formed with an athlete's name seems to *Jordan-esque* be at (102), with 48 hits, referring most often to Michael Jordan, who is a former basketball player and businessman. Examples of *Jordan-esque* seem to refer to other basketball players' achievements being almost as impressive as Michael Jordan's, or his specific moves on the basketball court, like in "I think the way Kawhi palms the ball is extremely Jordan-esque" (iWeb, nbadraft.net).

The majority of the hundred most frequent "last name+*esque*" adjectives appear to refer to artists; painters, film directors and musicians among others. This may pertain the nature of their work, which is either entirely or partially visual and audible and thus more tangible. Perhaps it is simpler to identify patterns in a painter's work or a writer's text, rather than a politician's speech or an athlete's movement. A visual artist might prefer certain colours or colour schemes, shapes, or have a unique visual style, which would be possible to emulate in a manner that justifies comparison. In a more extensive study, the names could be categorized more thoroughly, and it could be examined whether people writing about sports and athletes prefer to choose a different derivational suffix or another way of comparing a promising new athlete to a legendary one more often than using *-esque*.

4.2.2 Media proper noun formations

The second broad category observed is the media-related proper noun + *esque* adjectives. These media formations include films, television series, books, games, and characters in them, as well as two "media groups", the previously mentioned band *Beatles* and comedy group *Monty Python*. As

can be observed from the table below, the hyphenated version of the words is clearly preferred in this category. This may be affected by the fact that many of the examples are formed of a compound noun proper name. The preceding parts of the compounds are found in either the corpus' 'Context' tab or a separate query for the specific word. Only the more frequent *(Monty) Pythonesque* and the blended *Beatlesque* are found in the top hundred list, and they are the only two that refer to groups of people.

No.	Pos in 100	Word	Freq.	Vowel	Cons.	Additional Notes
1	17th	Bond-esque	212		C -	
2	18th	Pythonesque	193		C	
3	28th	Beatlesque	140		C	Blend
4	29th	Gatsby-esque	138	V -		
5	39th	Python-esque	98			Both versions = 291
6	41st	Beatles-esque	92			Both versions = 232
7	53rd	Zelda-esque	81	V -		
8	54th	Jones-esque	83		C -	50 Indiana, 9 Bridget
9	58th	Man-esque	74		C -	19 Iron, 9 Rain, 8 Mega, etc.
10	60th	Potter-esque	70		C -	
11	61st	Wonka-esque	70	V -		
12	64th	Frankenstein-esque	65		C -	
13	71st	Men-esque	61		C -	59 Mad
14	72nd	Runner-esque	61		C -	55 Blade
15	73rd	Matrix-esque	60		C -	
16	77th	Wars-esque	58		C -	49 Star
17	79th	Thrones-esque	57		C -	
18	89th	Diablo-esque	53	V -		
Most frequent without hyphen/with hyphen				-/4	2/10	

Table 2. iWeb Corpus "*esque" frequency results of media proper name+*esque/-esque* formations

Ten of these adjectives are formed of a name that is both the name of a film, game, book, series, or any combination of these, but also the name of a character in them. It is not always evident from the short context, whether the adjective describes something similar to the character or the style of the book or film. *Bond-esque*, seems to most often refer to something in the style of the films, like: "the first was the James Bond-esque missile base" (iWeb, kentonline.co.uk), but it is also used to describe something in the style of the character, like: "a picture of swooping lines,

elegant curves and classic James Bond-esque style" (iWeb, watchfinder.co.uk). Although *The Great Gatsby* by F. Scott Fitzgerald from 1925 is a popular book and *Jay Gatsby* its eponymous character, *Gatsby-esque* in the corpus examples mostly refers to the extravagant upper-class 1920's aesthetic of the 2013 film by Baz Luhrmann. The case is similar with the word *Potter-esque,* which mainly refers to the magical world of the *Harry Potter* franchise books and films, although there are references to "unfashionable Harry Potter-esque glasses" (iWeb, blogs.crikey.com.au) and "a killer Harry Potter-esque scar across his face" (iWeb, nzgamer.com). Interestingly, *Wonka-esque* (usually with preceding *Willy*) apparently most often refers to something in the style of the *Willy Wonka* company from Roald Dahl's novels, how the sweets in the films look or how the strange factory is realized. This is reminiscent of how *Disneyesque* does not refer to its originator but the work of the company. The other examples, , *Jones-esque* (mostly film franchise and main character *Indiana*; but *Bridget* in 9, books and films and their main character), *Runner-esque* (usually *Blade Runner,* film franchise and occupation of main character), *Zelda-esque* (shortened name of game franchise and character) and *Diablo-esque* (refers to the style of the video game, although the main villain is called *Diablo*) seem similar, as in, they mostly refer to the piece of media and not the character its' named after.

In contrast, *Frankenstein-esque* seems refer to something constructed horribly of disparate parts like the professor's monster, e.g. "collecting enough car parts that one day they will have a complete (yet Frankenstein-esque) car?" (iWeb, digitizer2000.com) and "The original Thing . . . is a giant outer world Frankenstein-esque creature who regenerates limbs" (iWeb, dailyturismo.com). The most interesting example of these character-named formations, however, is *Man-esque*, with 74 hits, that divide into several different adjectives. The most common is *Iron Man-esque* with 19 hits, referring to something like the comic book and film character *Iron Man* or the films created of the character. *Rain Man-esque* (9 hits) refers to a 1988 Oscar-winning film, especially to a character or person like the eponymous main character. There are also 8 *Mega Man-esque* (Japanese video game franchise) and likewise 8 *Burning Man-esque,* comparing something to the experimental festival

held in a Nevada desert. *Man-esque*'s plural counterpart, *Men-esque* revealed to refer to period drama series *Mad Men* 59 times, once to *All the President's Men* (1976 political drama thriller), and one of the results was irrelevant, something *old man-esque*.

4.2.3 Company and product name formations

The remaining results in the hundred most frequent proper noun + *esque* adjectives that were recognisable as adjectives pertain to company names or their products. Among the hundred top results, there were two adjectives formed of technology companies' names, (22) *Apple-esque* with 155 hits and (90) *Google-esque* at 53 hits, and one of a product: the iPhone, Apple's line of smartphones, with (99) *iPhone-esque* with 49 hits. Further, expanding to the 300 most frequent **esque* list, the next examples referring to products (that are not films, books, games etc.), arrive at (143) with *Twitter-esque* with 39 hits, (150) *Facebook-esque* with 37 hits, (190) *Snapchat-esque* with 31 hits, which all refer to social media companies or their sites and applications. The only other apparent ones in this category are (223) *Ikea-esque*, reminiscent of the Swedish furniture company's products with 28 hits, and (241) *Java-esque* with 24 relevant hits, referring to the programming language or its related software platform.

5. Conclusion

This thesis aimed to present and describe some of the most common adjectives formed of a proper noun and the derivational suffix -*esque* in the iWeb Corpus from 2017. As hypothesised, *Kafkaesque* was the most frequent relevant adjective found in the data, even though it is far from the earliest coinages with the suffix (Franz Kafka died in 1924). The hundred most frequent list includes some earlier adjectives, like *Rubenesque,* reminiscent of work of Peter Paul Rubens (16th century) and *Dantesque* (Dante Alighieri, 13th century). Most of the most common formations with

real people's names were made with artists' names of different fields, and a few referred to politicians. With these proper nouns, using the hyphenated version was slightly preferred when the noun ended in a consonant, but strongly preferred with a final vowel. The media-related formations showed that the hyphenated version was preferred, and that it differs whether the adjectives refer to the piece of media (e.g. film or book) with the name, or the eponymous character.

In this sort of analysis, the amount of cultural trivia knowledge that the person analysing the corpus data results possesses has a noticeable effect on the efficiency of the analysis. Recognising half of compound words as proper nouns (e.g. *Wars-esque*), or guessing who last names refer to was a substantial part of the analysis process. Having to read dozens of concordance lines or do Google-searches of most of the results would have been arduous. Nevertheless, it is completely plausible that I have missed some proper noun + *esque* adjectives even in the top hundred list if I did not correctly acknowledge them. Some examples which I initially regarded as proper noun formations, like *alien-esque* were ignored when the concordance lines revealed that most of the examples did not refer to the film franchise *Alien*, but something suggestive of a creature from outer space.

In the future, it would be interesting to investigate other corpora for the suffix, to examine whether the age of the text (even by decade) or the writer's occupation (e.g. journalist in news corpus) changes the most common adjectives. There would likely be a correlation between the cultural significance or popularity of the person or piece of media and the frequency of the *-esque* formation. For example, there are only 7 instances of *Biden-esque* and *Bidenesque* in the iWeb corpus, because the current United States president Joe Biden had only been a vice president at the time the corpus data was gathered. It is not possible, however, to comment on the change over time in this thesis, because of the synchronicity of the iWeb corpus. Further research could also be done even in the iWeb corpus for this suffix, with the help of the 'Collocates' search function. It would be intriguing to examine collocating words for specific adjectives, such as what do people most

often refer to as being *Trump-esque* or *Reaganesque*. Investigating the adjectival formations with similar derivational suffixes and a proper noun, such as *-like* or *-ish* compared to *-esque* would be another way to expand the study.

References

Davies, Mark. 2017. *The iWeb Corpus*. Available from english-corpora.org/iWeb/

English Oxford Living Dictionaries. 2019. Oxford University Press. Available from en.oxforddictionaries.com [Accessed April 2019]

Liu, W., & Shen, H. (2012). "A Corpus-based analysis of English suffix -esque". *Theory and Practice in Language Studies*, 2(4), 767-772.

Mattiello, Elisa. 2017. *Analogy in word-formation: A study of English neologisms and occasionalisms*. Berlin, Germany: De Gruyter Mouton.

Merriam-Webster Dictionary. 2023. Merriam-Webster, Incorporated. Available from merriam-webster.com/dictionary/

Moon, Rosamund. 2008. "Lexicography and linguistic creativity". *Lexikos*, (18), 131-153.

Oxford English Dictionary. 2023. Oxford University Press. Available from oed.com

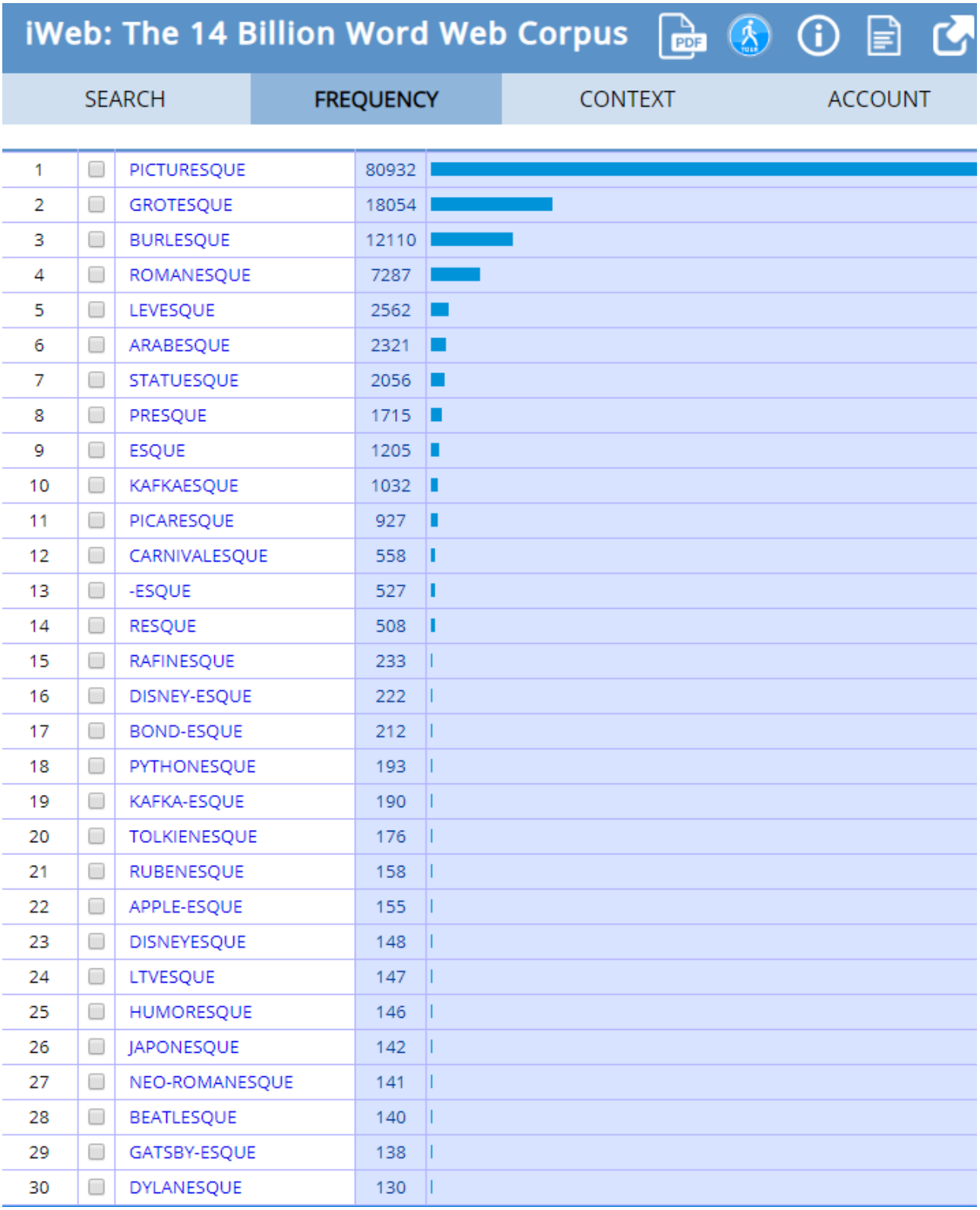
Plag, Ingo. 2003. *Word-formation in English.* Cambridge: Cambridge University Press.

Plag, Ingo, Maria Braun, Sabine Labbe, Mareile Schramm. 2009. *Introduction to English Linguistics*. Berlin: De Gruyter Mouton.

Weisser, Martin. 2015. *Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis*. Hoboken: Wiley.

Weisser, Martin. 2016/2020. *Corpus-based Linguistics.* Available from martinweisser.org/corpora_site/CBLLinks.html. [Accessed 24 April 2023]

Appendices



1		PICTURESQUE	80932
2		GROTESQUE	18054
3		BURLESQUE	12110
4		ROMANESQUE	7287
5		LEVESQUE	2562
6		ARABESQUE	2321
7		STATUESQUE	2056
8		PRESQUE	1715
9		ESQUE	1205
10		KAFKAESQUE	1032
11		PICARESQUE	927
12		CARNIVALESQUE	558
13		-ESQUE	527
14		RESQUE	508
15		RAFINESQUE	233
16		DISNEY-ESQUE	222
17		BOND-ESQUE	212
18		PYTHONESQUE	193
19		KAFKA-ESQUE	190
20		TOLKIENESQUE	176
21		RUBENESQUE	158
22		APPLE-ESQUE	155
23		DISNEYESQUE	148
24		LTVESQUE	147
25		HUMORESQUE	146
26		JAPONESQUE	142
27		NEO-ROMANESQUE	141
28		BEATLESQUE	140
29		GATSBY-ESQUE	138
30		DYLANESQUE	130

Appendix A. Top 30 Frequency List search for *esque

PAGE: << < 1 / 12 > >>

CLICK FOR MORE CONTEXT	[?]	SAVE LIST	CHOOSE LIST ---------	CREATE NEW LIST	[?]

1	hollywood-elsewhere.com	A	B	C	. Totally baffled by the love for the obnoxious, over directed, scream queens **esque** trash. The ESPN is so worth it though, but at 10 hours its
2	starmadedock.net	A	B	C	off small, and as you pick up materials etc you grow. The RTS **esque** elements are really a way to enhance that more then anything, you starting off
3	newbcomputerbuild.com	A	B	C	, which is- one of the best value cases out there and very Storm Trooper **esque**. The NZXT Phantom 240 has the following features and you can learn m
4	huskerboard.com	A	B	C	2010 that we thought was an intentional throwback to the old 70's Johnny Rodgers **esque** uniforms but actually weren't? The " super ultra mega extren
5	glamorous.com	A	B	C	Whether you stick with a classic biker style or opt for a more 70's **esque** look with fringe detailing, a suede biker jacket is this seasons answer to the
6	cavestory.org	A	B	C	and despite the older (non flixel, the engine that made the cave story **esque** game titled Condition) code it works pretty well. I even have boss and
7	clcboats.com	A	B	C	the possibilities of the strip-planking method. The result is this Aleut-like or Aleut " **esque** " design. Much has been written about the " bifid " or split bo
8	mmo-champion.com	A	B	C	the Doomhammer and using Windfury, whilst still retaining a lot of his Warcraft 3 **esque** Shaman abilities. # In World of Warcraft he's been very much l
9	blu-ray.com	A	B	C	, child genius Harold Wormser (Andrew Cassese) and very " Bluto " - **esque** Booger (Curtis Armstrong), find themselves powerless at Greek Council unti
10	nofilmschool.com	A	B	C	body of work has gone through several changes throughout his career, but the " **esque** " of his style really comes from his earlier work from the late 80
11	thesportster.com	A	B	C	man his size. His legendary brawls with HOFers Mick Foley and his Mortal Kombat **esque** entrance helmet earn Vaders time in the sun of the WWE Hall
12	nintendoeverything.com	A	B	C	a thingnext Zelda will be a eagle vision one, with multiplayer, open world **esque** type of game. # But heywhile everyone waits for Zelda, I am here
13	routeone.co.uk	A	B	C	and Guam the team got to sample some of the most creative and computer game **esque** spots around along with the amazing cultural delights that ca
14	geargods.net	A	B	C	up. This application is great for the chaotic hardcore stuff, like the HM-2 **esque** tones that are popular right now. Kurt Ballou is the guru on this subject
15	34st.com	A	B	C	with trendy vibes. The only caveats are the mildly tacky stock " photo " **esque** pictures and windows that open onto a relatively ugly parking lot. But if y
16	humblebeeandme.com	A	B	C	and I don't like the curls they give very much " very Shirley Temple **esque**, making them really hard to do much with. I should really put in
17	fadedspring.co.uk	A	B	C	The pint sized singer may have only been 5ft 1 but her classic pop Madonna **esque** vocals rang true and clear across House Festivals. Fans threw large
18	newparadigmastrology.com	A	B	C	Blessings and love!!! # Brother Kaypacha, Holy C = Very Zappa **esque**. Cradle to the grave we are all compost, maybe a rotten banana noir
19	metal-observer.com	A	B	C	rhythmic playing that's equal parts New York death metal density and Floridian, Morbid Angel **esque** occult tremolo texture. Rather than solely manifes
20	visitspacecoast.com	A	B	C	in Melbourne includes the Foosaner Art Museum, art galleries, the " Jaws " **esque** mural, and restaurants. Many locals can be spotted here throughout
21	4chandata.org	A	B	C	for the hell of it! " Even if it was some mindless r crumb **esque** shitshow with crackers killing niggers, why do these people feel like they ought to
22	freewillastrology.com	A	B	C	. Perhaps you weren't the butt of a cosmic joke or some Linda Blair- **esque** bad review, but in fact the recipient of a very precious gift of love
23	normanrecords.com	A	B	C	with a mix of NEU! motorik krautrock, Boards Of Canada ambience and Grimes **esque** pop. Dreamy and atmospheric but with a sense of urgency.view
24	tokunation.com	A	B	C	to Kamen Rider 1. # Super Hero Generations is a new Super Robot Wars **esque** tactical RPG from Bandai Namco Games that features Kamen Rider, Ult
25	steamforged.com	A	B	C	captain, scalpel is definitely closer to her. Although Scalpel is full of mortician **esque** tricks. a common sight is: # scalpel kicks off, leaves the ball
26	ilivehere.co.uk	A	B	C	alameida gardens does not count it is concrete and a burger king wrapper rolling tumbleweed **esque** down the street does not count # 2/Glacias and L
27	thenerdsofcolor.org	A	B	C	hinted this might be sort of like Hunger Games rest of the world a medieval **esque** landscape while this Astra or Asra? is an advanced city like the Capit
28	hopperhq.com	A	B	C	the use of filters. However these are not your average Lo-fi and Valencia " **esque** filters. The Modern Art Filters' that come preloaded with Prisma will m
29	weallsew.com	A	B	C	showed up as a Scottish Bagpiper. I love this shirt because its very Morticia **esque** but can be a base for any type of costume: witch, bat,
30	wesleying.org	A	B	C	37573300 6573300/ Tag Archives: Burlesque # **ESQUE**:: wtf is it??? Honestly, we don't know.
31	wesleying.org	A	B	C	. But were doing a show and we need YOU. # Last Spring, **ESQUE** took the form of a queer(ed) burlESQUE about desire- and difference. This April,