

Handsome Women, Gorgeous Men: A corpus-based analysis of frequency, animacy and gender over time in adjectives of beauty

Kit Callaway

University of Georgia – kec47019@uga.edu

Abstract

In 80,016 tokens of adjective-noun pairings, taken from historical (1810-2009) corpus data, 13 beauty adjective types were examined for changes in frequency patterns over time and to discover if their noun collocations were frequently masculine, feminine animate, or inanimate. The most common types are consistent in frequency, while less-common types are subject to trends. *Handsome*, *gorgeous*, and *good-looking* favor masculine nouns, and *pretty* favors feminine ones. *Good-looking* and *handsome* favor humans, and *lovely* and *beautiful* inanimates. These collocational meanings change over time. *Handsome* had a temporary peak of femininity in the late 19th century, while *gorgeous* has grown more animate.¹

1 Background

1.1 Introduction

There are many words to describe beauty. The Oxford Historical Thesaurus lists 69 total adjectives of beauty, from *beaut* to *wliti*. Obviously, some of these are more common than others and their popularity has changed over time. An adjective like *beauteous* is uncommon to see even in text now, but its frequency in the Corpus of Historical American English (Davies, 2010) shows that it was not always so unusual. While its frequency per million words was .07 in the 1980s-2000s sections of the corpus, in the 1810s it was 14.39 and it maintained a frequency of > 2.00 words per million throughout the 19th century. Tagliamonte and Brooke (2014) find that adjectives of strangeness have “trends” in usage, such that one form may be most popular in a certain era while another may be the most popular in other eras. Is this so in beauty adjectives also?

Likewise, the word *handsome* was once a common descriptor of well-formed faces, people, and things: *a handsome table* was once a perfectly ordinary noun

¹I would like to thank Dr. John Hale and Dr. Chad Howe of the University of Georgia for feedback and assistance on searching for and processing the data presented here.

phrase, whereas now it sounds less natural. In the 19th century *handsome* was particularly much-used in America, or so Martineau remarks in *Society In America*:

[Americans] use the word ‘handsome’ much more extensively than [the English] do: saying that Webster made a handsome speech in the Senate: that a lady talks handsomely, that a book sells handsomely. (Martineau, 1837, p. 210)

New England Magazine likewise adds in 1832, “Handsome is only applied more generally, as a term of commendation, than in polished English usage” (Buckingham & Buckingham, 1832). This is certainly at odds with the common use of *handsome* in American English now. Collocates in the Corpus of Contemporary American English (Davies, 2008) are strongly animate, and, beyond that, strongly masculine. The top five nominal collocates in the corpus’s word info for *handsome* are *man*, *face*, *boy*, *prince*, *feature*, and *devil*. Three of these are words with masculine meaning, and all are either human or parts of humans.

Intuitively, the space of beauty seems cordoned off in a way that is heavily responsive to gender and animacy as well as typical lexical change in word frequency over time. *That’s a pretty dresser* and *that’s a pretty woman* have a general pleasing-appearance reading, but *that’s a pretty man* intuitively means something more specific than just ‘good to look at’. *A good-looking man* and *a good-looking woman* are equally acceptable, but *a good-looking chair* is slightly odd. Findings on noun collocation (Moon, 2014; Norberg, 2016; Pearce, 2008) support the idea that words that are not grammatically gendered can nevertheless have gendered associations. Perhaps adjectives, too, can have associative meanings, with some being preferred for masculine or feminine nouns and some being preferred for inanimate nouns.

Yet these meanings, it seems, are not fixed: they change even over a relatively short span of time. To return to our earlier example, in the 1950s section of The Corpus of Historical American English (Davies, 2010) the phrase *handsome woman* has a frequency per million of 1.91, while 20 years later in the 1970s it drops to .50. (*Handsome man*, by contrast, is 1.35 in the 50s and 2.27 in the 70s). Thus it seems that it was ‘trendier’ to describe women as handsome in the 50s, at least in the kind of written publications that COHA captures, than it was in the 70s. Yet in both eras the word remained more common for describing men.

The objectives of this study are as follows:

1. Examine whether adjectives of beauty vary in frequency due to recycling, as Tagliamonte and Brooke (2014) suggests strangeness adjectives do.
2. Examine whether adjectives of beauty vary in their frequency depending on the animacy or gender of the noun they describe, as previous collocational research has suggested they may.

1.2 Literature Review

Recycling and renewal are very common in some parts of speech. They have been particularly well-studied in intensifiers. Articles such as Ito and Tagliamonte (2003) and (Xiao & Tao, 2007) describe the process as follows. An adverb is used to describe an adjective in a way that makes its meaning more vivid. But as the adverb becomes established in this role, it loses its literal meaning, and simply has an intensifying function. The power of the intensifier is partly lost due to this fading of the literal meaning. Thus new intensifiers are used to create more powerful meanings, and in their turn become grammaticalized, lose their intensity, and get replaced.

As Ito and Tagliamonte (2003) describe it, there is a constant process of fading, replacement and recycling happening in the intensifier space. Earlier forms coexist with later forms, but there is great variation in which intensifiers are most *frequent* in speech. The ‘popular’ forms, in effect, go in and out of style with time, even though the set of intensifiers does not completely change. Moreover, these ‘trends’ vary among social groups. Thus, for instance, in Ito and Tagliamonte’s study, speakers 66+ used *very* more frequently than *really* while speakers 17-34 used *really* more frequently than *very* (p. 267).

Can adjectives undergo a similar fading-and-recycling or ‘trendiness’? In at least one semantic domain this seems to be an active process, as Tagliamonte and Brooke (2014) found in their study of adjectives of strangeness. They discovered that the eponymous adjective for this space, *strange*, was rapidly being replaced by *weird* as the most common expression of this field. As with intensifiers, less-common adjectives are not being totally lost in this process of replacement. Rather, they are simply being upstaged in popularity.

The field of strangeness-adjectives underwent lexical replacement leading to layering, with the items falling out of vogue and being recycled as ‘fashionable’ over time (p. 13). What is noteworthy is that for the most part, new words are not being rapidly created and lost. Many of the words in the study are first attested before the 1700s. Instead the words change their frequencies and become more or less popular at a given time, without totally falling out of use.

The authors analogize this process directly to intensifier change, suggesting that, like intensifiers, adjectives can undergo renewal (whereby new forms enter the language and come to describe the same space as existing forms) and recycling (when related forms change in relative frequency over time, displaying ‘trends’) (p. 17). Tagliamonte and Brooke suggest that this is particularly the case with subjective adjectives that have strong positive or negative meanings (p. 10). If this analysis is correct, the process should be visible in other adjective spaces, not just one.

Beauty adjectives fit the criteria of being subjective and expressive. They have a strong positive meaning by nature of their semantic domain. Moreover, they have, as previous stated, gender and animacy associations that also seem to have the potential to change over time.

Other patterns of lexical change also exist. Fruehwald (2017) coins the

term *zeitgeist effect* to describe a type of change in which a whole community engages in a particular change at one moment in time. This change is tied to a time period rather than a generation or an age. It may then revert or go in a different direction. In the study in which he proposes this type of change, male speakers displayed strongly conservative pronunciation of /ow/, but only during the 1980s. Before and after the 1980s, they conformed more generally to the fronting of /ow/ that was going on in the rest of the population.

This is similar to renewal & recycling. But whereas renewal & recycling is dependent on demographic features like age, the *zeitgeist* effect occurs across a particular swathe of the population regardless of age. Fruehwald describes this in the context of changes to the vowel space, but it could equally apply to other linguistic dimensions. In the case of lexical change, a *zeitgeist* effect might involve temporary changes to the meaning or popularity of a word, so that it is used more frequently or in different contexts during a particular time span.

Both of these patterns are in contrast to traditional lexical change. Variationist sociolinguistics has typically understood lexical change as follows: new forms enter the language and is diffused throughout the language, such that it forms an S-curve. The form begins as a low-frequency item, gains frequency throughout time, and then becomes high-frequency (at least relative to its context) (Denison, 2003). Forms may also drop in frequency and be lost.

For a lexical set, then, there are three potential patterns we could expect to see a form follow over a period of change. It might become more frequent and stay frequent or become less frequent and stay less frequent, as in traditional lexical change. It might go in and out of popularity at a slow rate, as it is recycled and renewed, as per Tagliamonte and Brooke (2014). It might experience a fairly consistent frequency (low or high) except for a period where it spikes, as per the *zeitgeist* effect (Fruehwald, 2017). Items in the set of adjectives of beauty can be examined for each of these three patterns.

But it is difficult to deal with adjectives of beauty without also discussing the issue of gender, because the existing body of collocational research already suggests that there is some disparity in how they are applied to men and women. A growing body of research (Garg et al., 2018; Moon, 2014; Norberg, 2016; Pearce, 2008) supports the idea that a word can have associations not directly present in its central meaning, as supported by collocations of multiple words. Social meanings such as gender and ethnic stereotypes are two areas where collocation research has revealed loose associations between particular words. As Romaine (2001) puts it, “Collocations transmit cultural meanings and stereotypes which have built up over time” (p. 109).

A good deal of gendered-collocation research has, quite reasonably, begun with a nominal item, like *man* or *woman*, and investigated the other words that collocate with it. Moon (2014) identifies adjective collocates most common with men and women of different age groups. The most common collocates for young woman are *beautiful*, *attractive*, and *pretty*, whereas for young man they are *handsome*, *nice*, *bright*, and *tall* (pp. 15-17). While beauty adjectives are very frequent for women in this study, they are much less frequent for men.

The adjective *sexy* was also far more frequent for young women than it was for young men (p. 17). This supports the idea that pleasing-appearance adjectives are distributed differently when the noun they are modifying is masculine vs feminine.

Similar results occur in Norberg (2016). *Handsome* collocated far more often with *boy* and *beautiful*, *lovely* and *pretty* with *girl*, but beauty adjectives generally favored *girl* more than they favored *boy* (p. 305). *Beautiful* is the top adjective collocate for *girl* in this study; *lovely* is fourth, and *pretty* is fifth. By contrast, for *boy*, *handsome* comes in seventh place, and it is the only pleasing-appearance adjective to make the top ten collocates.

Pearce (2008) reinforces again that even when men are described with personal beauty adjectives, the adjective choice differs. In this dataset, women's pleasing appearance was most often described with *attractive*, *beautiful*, and *pretty*, with the latter referring exclusively to women. Men's pleasing appearance was described as *best-looking*, *fantastic-looking*, *devastating-looking* (p. 17). Although it is possible to describe women as *handsome* and men as *beautiful*, the reverse is much more common.

Clearly, then, this is semantic domain in which adjectives behave in gendered ways. The adjectives that we select to describe beauty in men are not necessarily the same as those we select to describe beauty in women, even allowing for the fact that women are described with beauty adjectives far more often than men.

Animacy in adjectives of pleasing appearance is less explored through collocational research, but intuition suggests that it may be relevant. *Handsome's* top nominal collocates in COCA are man, face, boy, prince, feature, devil, and fellow; these are all animate items, with the possible exception of face and feature, which do apply largely to animate nouns. *Beautiful's* top collocates are woman, girl, god, flower, dress, America, creature, and landscape – flower, dress, America, and landscape, four of the six, are inanimate. With this disparity, it seemed wise to take into account animacy when designing this study.

2 Methods

The corpus used was The Corpus of Historical American English (Davies, 2010). This four-million-word corpus contains fiction, nonfiction books, popular magazines, and newspapers from 1810 to 2009. Although the number of words per decade are not balanced, genre percentages are balanced decade-by-decade, so that the percentage of (for example) fiction is the same in each decade. The corpus is annotated by decade and by genre. The search was conducted on the UGA Corpus Server, which was built using the IMS Open Corpus Workbench (Evert, 2008).

As previously mentioned, the area of investigation is *adjectives of beauty*. I have chosen this category because a) it is relatively common in everyday speech and writing (often *beautiful* and *handsome* are in the top collocates for *man* and *woman*, for example); b) individual words within the set often have strong

gender/animacy intuitions; c) at least one item in the set (*handsome*) seems to have been ‘trendy’ and then fallen somewhat out of fashion, suggesting some degree of active nonlinear change ongoing in the space.

The most basic item of the beauty set is obviously *beautiful*. The Oxford Historical Thesaurus Online was used to locate other historical adjectives of pleasing appearance. Synonyms for *beautiful* were also sought in several modern thesauruses. Of these items, only words which described a general characteristic were selected, not one specific to a given body part (like *bright-eyed*). Words which did not have at least 50 attestations clearly referring to beauty in COHA were also removed. The search targeted only attributive adjectives, ignoring predicative ones. This was done to narrow the set and standardize the context in which each adjective was appearing. Attributive adjectives were chosen because they were closer to their corresponding noun, so it was easier to be sure of the referent. Thus, all adjective preceded a noun in an NP. The ensuing dataset consisted of 80,016 total tokens of 13 types: *beautiful*, *handsome*, *attractive*, *pretty*, *lovely*, *beauteous*, *cute*, *sexy*, *gorgeous*, *comely*, *bonny*, *good-looking*, and *ravishing*.

Table 1: Token Counts by Decade and Type

Decade	Beauty Adjectives	Total Words	Type	Tokens
1810s	221	1,181,022	beautiful	32952
1820s	1587	6,927,005	pretty	14734
1830s	3825	13,773,987	lovely	11171
1840s	4878	16,046,854	handsome	10398
1850s	4902	16,493,826	attractive	4329
1860s	4135	17,125,102	gorgeous	2678
1880s	5669	20,872,855	good-looking	966
1890s	5716	21,183,383	cute	757
1900s	5368	22,541,232	beauteous	552
1910s	4591	22,655,252	sexy	570
1920s	4781	25,632,411	comely	520
1930s	3771	24,413,247	bonny	219
1940s	3552	24,144,478	ravishing	169
1950s	3594	24,398,180		
1960s	3293	23,927,982		
1970s	3294	23,769,305		
1980s	3377	25,178,952		
1990s	3791	27,877,340		
2000s	4603	29,479,451		

Frequencies were normalized using the number of attributive adjectives per era in the corpus. This was done for two reasons. Firstly, the extremely low-frequency nature of many of these adjectives meant that they had a rounded frequency of 0 when considered against the millions of words in each decade. Secondly, stylistic considerations mean that number of total attributive adjectives might differ decade to decade. When frequencies are given for the types, they are considered as a percentage of **total attributive adjective use per time period**, not percentage of total words per time period. Because by-decade analysis proved statistically unwieldy given the number of decades, the data was split up into approximately 30-year eras as follows: Early 19th (1810-1839); Mid 19th (1840-1869); Late 19th (1870-1899); Early 20th (1900-1929); Mid 20th (1930-1959); Late 20th (1960-1989); and Early 21st (1990-2009). Although this last era is a decade short, the highest concentration of words come from this period.

Separate logistical regressions using lme4 (Bates et al., 2015) were conducted using animate/inanimate as the response value on the full dataset and using masculine/feminine as the response value in the gendered subset of the data. Two case study adjectives were also given separate regression analyses. Genre was treated as a random factor; text and noun were tested but had too many levels and induced technical problems. Since texts were divided by genre, genre proved a less unwieldy proxy. Graphs were created using ggplot2 (Wickham, 2016).

3 Results

3.1 Frequency

An unexpected item was the distribution of adjectives of beauty throughout the timespan. Numbers of beauty adjectives were highest in the late 19th century, even though this era did not have particularly high word counts compared to the 20th and 21st century, nor did it have more attributive adjectives total. Figure 1 shows normalized frequency of beauty adjectives (beauty adjectives per era over all attributive adjectives per era).

The peak is in the late 19th century (1870-1899) and the low point in the late 20th (1960-1989). This is probably a reflection of stylistic preferences in the writing of the periods. Note that this is normalized frequency of beauty adjectives relative to all attributive adjectives, so if it does reflect stylistic concerns, it is an increase in the tendency to describe things as beautiful that is visible here, rather than an increase in the absolute tendency to use attributive adjectives. Though outside the scope of the present study, this may be fruitful ground for a stylistics analysis examining the effects of various literary movements, such as the Romantics and the Modernists, on beauty adjective use. That we are seeing a recovery in the early 21st century may indicate a stylistic shift in progress.

As one might expect, the 13 types were not distributed exactly equally

Figure 1: Bar graph of beauty adjectives by era, normalized by number of attributive adjectives per era.

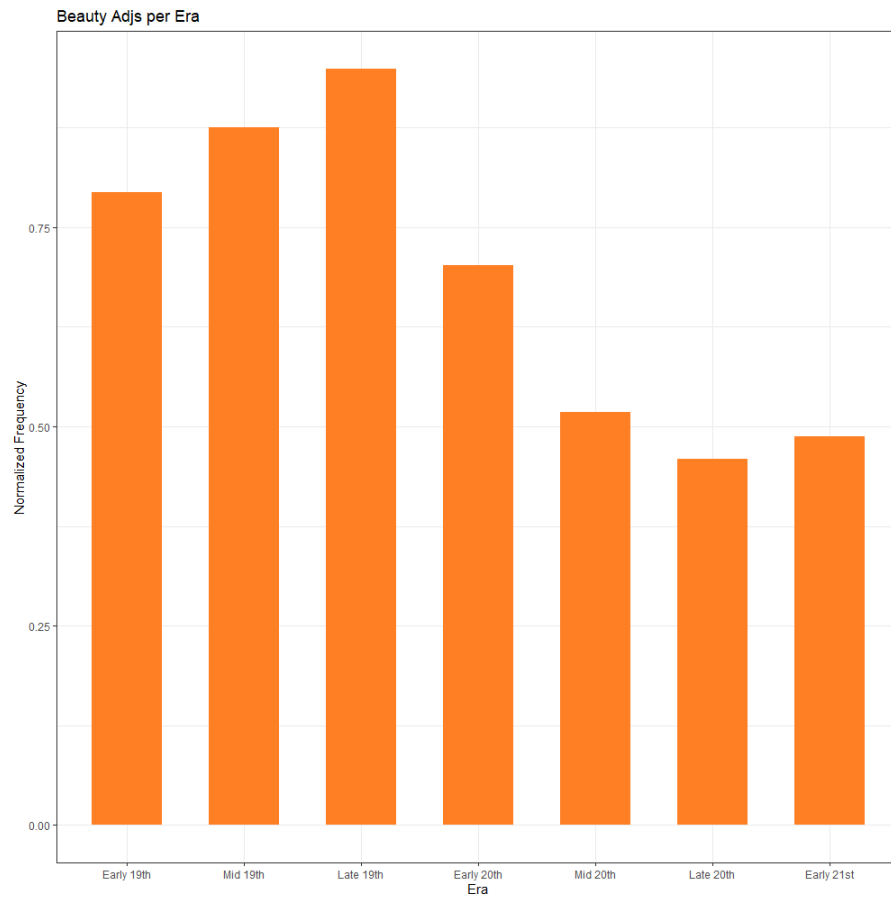


Figure 2: Bubble chart of adjective frequencies over time. Dot size represents normalized frequency.

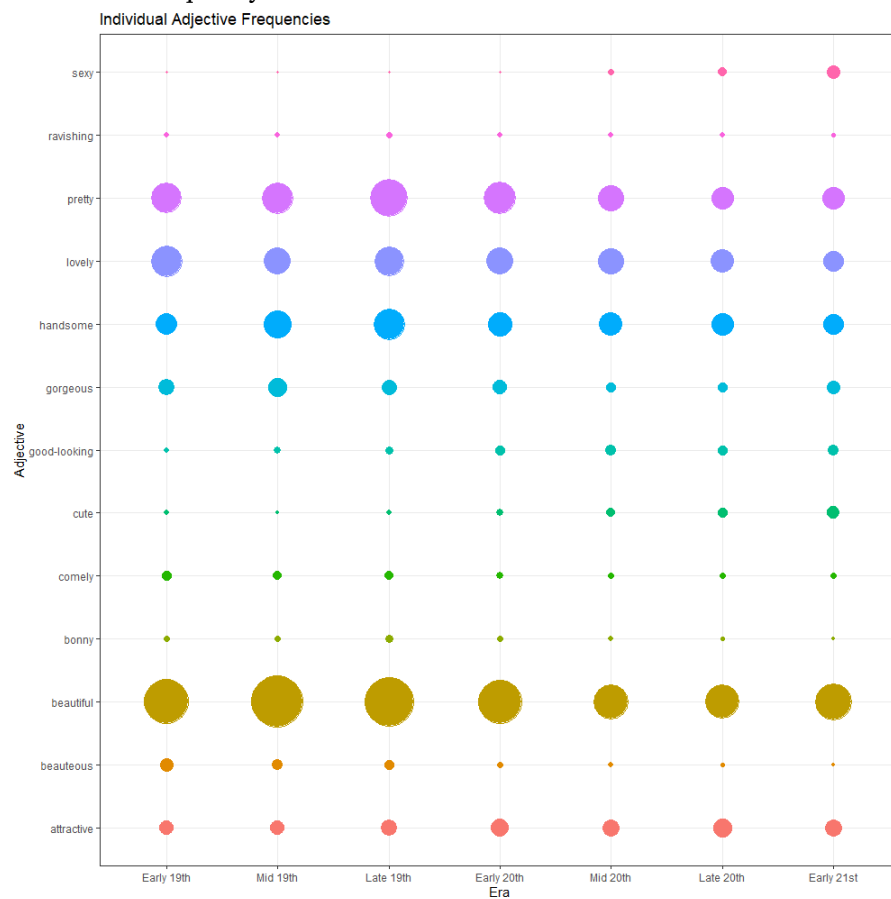


Table 3: Normalized frequencies of adjectives by decade.

	Early 19th	Mid 19th	Late 19th	Early 20th	Mid 20th	Late 20th	Early 21st
beautiful	0.305	0.417	0.382	0.292	0.183	0.167	0.195
pretty	0.134	0.146	0.213	0.152	0.097	0.069	0.068
lovely	0.141	0.102	0.127	0.098	0.092	0.071	0.057
handsome	0.063	0.112	0.140	0.081	0.075	0.066	0.055
attractive	0.023	0.025	0.032	0.038	0.037	0.046	0.037
gorgeous	0.032	0.043	0.028	0.023	0.009	0.009	0.022
sexy	0.000	0.000	0.000	0.000	0.002	0.007	0.021
cute	0.002	0.000	0.001	0.003	0.008	0.009	0.017
good- looking	0.001	0.004	0.005	0.008	0.010	0.009	0.011
comely	0.009	0.007	0.006	0.004	0.003	0.002	0.002
ravishing	0.001	0.002	0.002	0.001	0.001	0.001	0.001
bonny	0.002	0.002	0.005	0.002	0.001	0.001	0.000
beauteous	0.020	0.012	0.008	0.002	0.001	0.001	0.000

across periods. Some are more popular early on and grow rarer, some come into existence and increase in frequency. As this happens, the semantic space is jostled, with the space each individual adjective takes up changing over time.

Unsurprisingly, *beautiful* is the most frequent option of the 13 types across all eras and decades. For the most part, the other high-frequency adjectives are consistent. The general ranking, in order of frequency, is *beautiful*, *pretty*, *lovely*, *handsome*, *attractive*, *gorgeous*. There is some competition between *pretty* and *lovely* and between *attractive* and *gorgeous*. As might be expected, *handsome* experiences a popularity spike in the mid to late 19th centuries, which moves it from fourth to third place. Between the early 19th century and the mid 19th century, it almost doubles in frequency, and between the late 19th and early 20th it almost halves. This is consistent with a zeitgeist effect favoring *handsome*.

The remaining seven are where serious changes to frequency over time took place. *Beauteous* and *comely*, while overall low-frequency adjectives, were much more popular in the early to mid-19th century than they are now. For instance, *beauteous* was 8th place in the early 19th century whereas now it is tied for last place with a rounded frequency of 0. *Good-looking*, *cute*, and *sexy* became more popular over time. *Bonny* experienced a minor peak of popularity in the late 19th century. *Gorgeous* dipped sharply in popularity in the mid-late 20th century before recovering. *Ravishing* stays consistently low over time.

Some of these changes have obvious cultural motivations. For instance,

sexy arises in the early 20th century (first appearing in this corpus in the 1920s) because the taboo around discussing sexual topics directly was lessening. But most of them are apparently motiveless. Why did *beauteous* fall out of favor? Why is *comely* now less common than *cute*, whereas in the early 20th century they were roughly equal?

It seems that unlike Tagliamonte and Brooke's (2014) results for strangeness, the most frequent items in the pleasing appearance set do not significantly shift in relative frequency over time. This would seem to align with more typical lexical change – where items become overall more or less frequent over time and stay that way – or with the zeitgeist effect, rather than with recycling & renewal.

3.2 Gender & Animacy

For animacy, distribution slightly favored inanimates. The dataset was 59% inanimate and 41% human. The adjectives most strongly biased towards humans were *good-looking* (odds ratio 0.17, $p < .01$) and *handsome* (odds ratio .50, $p < .01$). *Lovely* (odds ratio 1.93, $p < .01$) and *beautiful* (odds ratio 1.46, $p < .01$), as well as the 'other' category containing the least high-frequency verbs (odds ratio 1.45, $p < .01$), were most strongly biased towards inanimates. In terms of era, the further back one goes, the less strongly animate the adjective use overall gets. This suggests either that the chosen adjectives have become more personal or animate over time, or that we describe inanimate objects with beauty terms less often now. Again, stylistic considerations in the texts involved may be at play. The early 20th century (1900-1930) represents the point where this trend becomes statistically significant.

Only 24% of the ~80,000 tokens in the dataset were nouns with gender. In these there was a notable preference for using adjectives of beauty to describe women rather than men. Of the 19,332 gendered tokens in the dataset, 14,734 (76%) were used for women, while only 4,598 (24%) were used for men. This is unsurprising as it is in line with the findings of other collocational gender analyses mentioned in the literature review.

The most masculine adjectives of beauty are *handsome* (odds ratio 40.31, $p < .01$), *good-looking* (odds ratio 11.09, $p < .01$), and *gorgeous* (odds ratio 4.22, $p < .01$). The most feminine adjective of beauty is *pretty* (odds ratio .36, $p < .01$). The others are, by and large, equal. What this probably really means is that all of those which do not attain statistical significance are equally strongly biased towards women, given the high ratio of female to male tokens in the gendered dataset. Overall, then, most of the adjectives on the list are preferred for women over men, and only *good-looking*, *handsome*, and, to a lesser degree, *gorgeous*, are more likely to be used with men, while *pretty* is even more strongly feminine than the others.

One may note that *good-looking* and *handsome* most strongly favor animacy and most strongly favor masculinity. In terms of what adjective can be used to describe them, it seems that inanimate objects and women pattern more closely together than either does with men.

Table 5: Distribution of adjectives used to describe feminine, masculine, and inanimate objects. X-squared = 123.37, df = 14, p-value < 2.2e-16.

	attractive	beautiful	good-looking	gorgeous	handsome	lovely	pretty	other	sum
feminine	5.44	38.84	1.81	0.79	6.64	15.68	27.11	3.69	100
masculine	5.39	12.59	10.72	1.91	55.92	3.41	5.74	4.31	100
inanimate	5.77	46.17	0.19	4.87	9.56	13.56	16.76	3.12	100

The adjectives for women, in order of frequency, are *beautiful*, *pretty*, *lovely*, *handsome*, *attractive*, other, *gorgeous*, and *good-looking*. The most common adjective for inanimate objects are *beautiful*, *pretty*, *lovely*, *handsome*, *attractive*, *gorgeous*, other, and *good-looking*. These lists differ only when one reaches the single-digit percentages. By contrast, for men the most common adjective is *handsome* (overwhelmingly), followed by *beautiful* and then *good-looking*, with *pretty*, *attractive*, *lovely*, other, then *gorgeous*. One may easily describe a woman or a dresser as *beautiful* or *pretty* but less easily as *handsome* or *good-looking*. By contrast, one may describe a man as *handsome* or perhaps *good-looking* or *beautiful* but not *pretty* or *lovely*.

One is wary of making too much soup from one oyster. Nevertheless there seems to be some sort of unmarked beauty-word territory, the words in which may be applied freely to women or to objects but less freely to men. As previously discussed, past literature has suggested that men are less frequently described in beauty terms compared to women. It may be that men are less likely to be described as *beautiful* or *pretty* because their aesthetic evaluability is not as salient as either women's or the types of inanimate objects that tend to occur with beauty adjectives.

3.3 Case Studies: Individual Adjectives

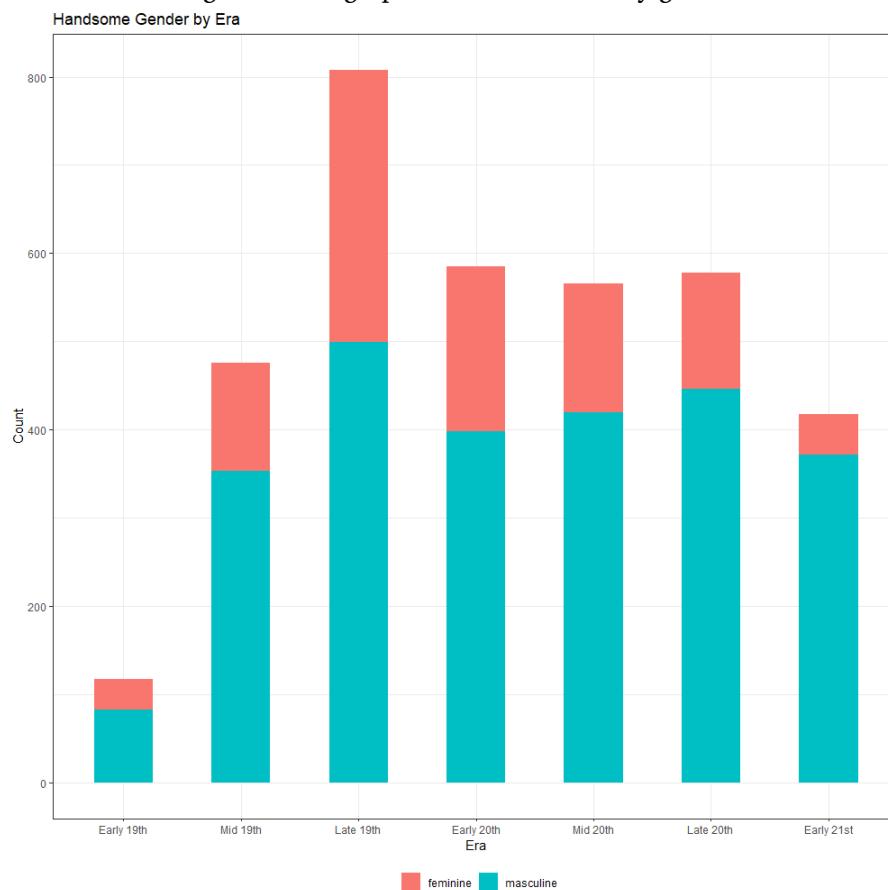
Gender and animacy associations have fluctuated over time for many of these adjectives but not necessarily in consistent ways. The two adjectives that have the most interesting diachronic story are *handsome* and *gorgeous*.

Handsome has had a varied career. As discussed, it experienced a spike in use in the late 19th century. The normalized frequency of *handsome* in the 1830s is .08 and in the 1910s it is .077, but at its peak in the 1880s it is 0.16, beating out *lovely* for third place. This change might well be described as a temporary trend, since it did not become entrenched. Collocational meanings also changed over time. Its proportion of collocating feminine nouns in the early 19th and early 21st century is similar, but in the five eras between those times it experienced a general rise, followed by a decline, in the proportion of feminines versus masculines. *Handsome* is the least masculine in the late 19th century.

Table 7: Handsome odds ratios by era and gender. Reference level is feminine.

<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.20	0.13-0.29	< 0.001
Late 20th	0.32	0.18-0.56	< 0.001
Mid 20th	0.18	0.10-0.33	< 0.001
Early 20th	0.11	0.06-0.21	< 0.001
Late 19th	0.10	0.05-0.21	< 0.001
Mid 19th	0.27	0.08-0.89	0.032
Early 19th	0.23	0.03-1.80	0.161

Figure 3: Bar graph of handsome era by gender.



Meanwhile, *gorgeous* has been getting increasingly animate as it loses its meaning of ‘showy, splendidous’ and simply comes to mean ‘beautiful’. In noun terms, earlier uses of *gorgeous* occur quite frequently with clothing, textiles, and art. In later usage, particularly in the 1960s onward, it becomes increasingly used with people.

Table 8: Gorgeous odds ratio animacy by era. Reference level is human.

<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.51	0.89-2.55	0.126
Late 20th	1.40	0.93-2.11	0.107
Mid 20th	4.92	2.95-8.20	< 0.001
Early 20th	5.23	3.31-8.25	< 0.001
Late 19th	9.74	5.61-16.93	< 0.001
Mid 19th	6.82	3.93-11.84	< 0.001
Early 19th	8.19	3.34-20.06	< 0.001

Gorgeous and *handsome* are interesting cases of contrast because they illustrate two different ways in which language can vary over time. In *gorgeous* we have a very classic case of semantic extension. Because *showily beautiful* makes very little sense as a description of a person in most contexts, it is very infrequently applied to animate subjects early in its career. But as the meaning becomes less narrow, it becomes more and more common to use it with people. This generational change intensifies over time and has thus far remained directionally consistent.

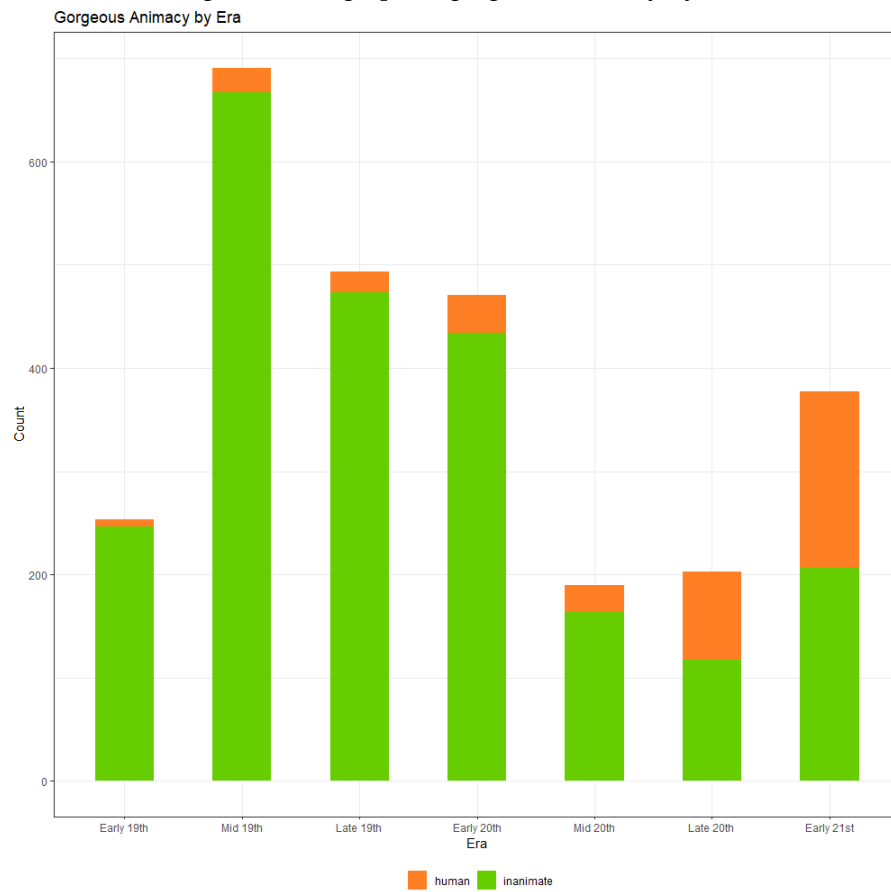
On the other hand, *handsome* is less linear. It had a peak use around the late 19th century. In this period, it was the third-most-popular word of this dataset to use in aesthetic judgements. One might expect it to follow the same path as *gorgeous*, becoming increasingly generic as time goes on and staying that way. Instead, it gradually drops back down in frequency and masculinity to its former position. In other words, there was a fad for *handsome* as a general, unmarked beauty word which then subsided. This is more akin to what Fruehwald (2017) calls the zeitgeist effect.

3.4 Conclusions

Change in lexical frequency could be described with three potential paths:

1. Typical lexical change, in which items become more popular or less popular over time with less popular items eventually becoming lost.
2. Recycling and renewal, where popular items in a limited lexical set jostle with each other for popularity over generational timespans creating layering.

Figure 4: Bar graph of gorgeous animacy by era.



3. The Zeitgeist effect, where a new or existing word becomes changes frequency or meaning for a short period of time and then returns to its previous status.

In this dataset 1 and 3 are present for adjectives of beauty, but 2 is not. The highest-frequency items generally remain static ranks in the corpus. This is especially so for *beautiful*, which remains consistently the most adjective common in the corpus to describe pleasing appearance. This fails to replicate the findings of Tagliamonte and Brooke (2014), suggesting that adjectives of pleasing appearance are not subject to the same types of change as adjectives of strangeness.

As one might expect with any large set of lexical items, low-frequency items sometimes become less common or are lost entirely, and new items may enter the set. However, these same low-frequency items may experience temporary, apparently unmotivated changes in usage for brief periods. This is the Zeitgeist effect. What is still needed is a better way to quantify that popularity. A statistical measure for testing the significance of normalized frequencies (rather than means or proportions) would be necessary to carry this out properly. Whether the adjectives' popularity differs among various social groups is another issue that the data from this study cannot address.

In this dataset, beauty adjectives respond to gender and animacy in nouns. Which adjective is selected to describe something beautiful depends on whether the noun is human, and if it is human, whether it is a man or a woman. Women are described overall more with beauty adjectives than men; this is in line with previous gender-collocation research. Men are described as *handsome* and *good-looking* more often. The other high-frequency beauty adjectives can be used to describe both women and inanimate objects with relatively equal ease.

Animacy and gender collocation is not static over time. Like frequency, it is subject to trends in specific words. Handsome experienced a phase of being less strongly tied to animate, male nouns in the late 19th century. Gorgeous has experienced a growth of animacy over time that is similar to growth of frequency over time. As such, collocational meanings in lexical items can be understood as being subject to the same patterns of change as frequency is.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Buckingham, E., & Buckingham, J. (1832). *Yankeeisms* (Vol. 3).
- Davies, M. (2008). The corpus of contemporary american english (coca): 560 million words, 1990-present. *Corpus of Contemporary American English (COCA)*, 24, 2018. <https://corpus.byu.edu/coca/>

- Davies, M. (2010). The corpus of historical American english (coha): 400 million words. 24(2019), 1810–2009. <https://www.english-corpora.org/coha/>
- Denison, D. (2003). Log(ist)ic and simplistic s-curves. In R. Hickey (Ed.), *Motives for language change* (pp. 54–70). Cambridge University Press. <https://doi.org/10.1017/CBO9780511486937.005>
- Evert, S. (2008). The ims open corpus workbench. <http://cwb.sourceforge.net/index.php>
- Fruehwald, J. (2017). Generations, lifespans, and the zeitgeist. *Language Variation and Change*, 29(1), 1–27. <https://doi.org/10.1017/S0954394517000060>
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16), E3635–E3644. <https://doi.org/10.1073/pnas.1720347115>
- Ito, R., & Tagliamonte, S. (2003). Well weird, right dodgy, very strange, really cool: Layering and recycling in english intensifiers. *Language in Society*, 32(2), 257–279.
- Martineau, H. (1837). *Society in america* (Vol. 2). Project Gutenberg. <https://www.gutenberg.org/files/52685/52685-h/52685-h.htm>
- Moon, R. (2014). From gorgeous to grumpy: Adjectives, age and gender. *Gender and Language*, 8(1), 5–41. <https://doi.org/10.1558/genl.v8i1.5>
- Norberg, C. (2016). Naughty boys and sexy girls: The representation of young individuals in a web-based corpus of english. *Journal of English Linguistics*, 44(4), 291–317. <https://doi.org/10.1177/0075424216665672>
- Pearce, M. (2008). Investigating the collocational behaviour of man and woman in the bnc using sketch engine. *Corpora*, 3(1), 1–29. <https://doi.org/10.3366/E174950320800004X>
- Romaine, S. (2001). A corpus-based view of gender in british and american english. *Gender Across Languages*, 153–175.
- Tagliamonte, S. A., & Brooke, J. (2014). A weird (language) tale: Variation and change in the adjectives of strangeness. *American Speech*, 89(1), 4–41. <https://doi.org/10.1215/00031283-2726386>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org/>
- Xiao, R. Z., & Tao, H. (2007). A corpus-based sociolinguistic study of amplifiers in british english. *Sociolinguistic Studies*, 1(2), 3055. <https://doi.org/10.1558/sols.v1i2.241>