



A study of the semantic preference and semantic prosody associated with the denominations of aging people

Denis Jamet, Christophe Coupé

► To cite this version:

Denis Jamet, Christophe Coupé. A study of the semantic preference and semantic prosody associated with the denominations of aging people. *Journal of Language and Aging Research*, 2024, 2 (1), <https://journals.sub.uni-hamburg.de/hup2/jlar/article/view/1183>. 10.15460/jlar.2024.2.1.1183 . hal-04535195

HAL Id: hal-04535195

<https://hal.science/hal-04535195>

Submitted on 6 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

A study of the semantic preference and semantic prosody associated with the denominations of aging people

Denis Jamet^{1*} & Christophe Coupé²

¹Jean Moulin University, France

²CY Cergy Paris University, France

Abstract. This paper studies the representation and conceptualization of aging people depending on the way they are referred to through a series of near synonyms: *old* / *older* / *elderly* / *middle-aged* / *mature* + N. A specific focus will be laid on the notions of “semantic preference,” i.e., “the tendency for an item to co-occur with a set of semantically related words,” and “semantic prosody,” a.k.a. “evaluative, pragmatic, emotive or attitudinal prosody,” to account for the evaluative aspect linked to certain denominations. This paper compares semantic preference and semantic prosody with five near-synonymous adjectives. It shows that denominations of aging people acquire a specific evaluation—a negative or positive semantic prosody—through their semantic preference, in that specific case the different attributive adjectives used in prenominal position of the sequences under scrutiny. This study addresses the question of the frequent negative axiology linked to the discourses on old age and aging through the notion of semantic prosody and will mostly examine the potential gender differences in the denominations of aging people.

Keywords. old age, aging, corpus linguistics, semantic preference, semantic prosody, evaluation, representations

A peer-reviewed contribution to *Journal for Language and Aging Research* (JLAR).

Submitted: 2023-05-10

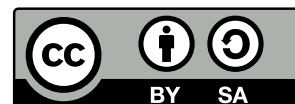
Accepted: 2024-01-04

Published: 2024-04-06

DOI: 10.15460/jlar.2024.2.1.1183

© Denis Jamet & Christophe Coupé

This work is licensed under a **Creative Commons** “Attribution-ShareAlike 4.0 International” license.



1 Introduction

The notions of “old age,” “aging,” and their representations have been brought to the front in recent years, especially so with the Covid-19 pandemic which has questioned

*. Corresponding author, denis.jamet@univ-lyon3.fr



Please cite as Jamet, Denis and Christophe Coupé. 2024. “A study of the semantic preference and semantic prosody associated with the denominations of aging people.” *Journal of Language and Aging Research* 2(1): 47–84. 10.15460/JLAR.2024.2.1.1183.

our views of and on the elderly. This article intends to focus on the evaluative aspect of the gendered denominations of aging people, through analyses carried out on the English Web 2020 corpus (available on SketchEngine 2017–2024, see Kilgarrieff et al. 2014). We chose to refer to *old* / *older* / *elderly*, etc. people by the phrase “aging people” (AP from now on). The notions of “semantic preference” and “semantic prosody”² are used to assess the evaluation triggered by the use of specific axiological adjectives used in combination with the denominations for AP. Taking this into account, the current study is structured around the following questions:

- Are the denominations of AP neutral labels attached to a referent, or are they evaluative depending on the word choice?
- Do the adjectives used in combination with said denominations (semantic preference) have an effect on their evaluative aspect (semantic prosody)?
- Can we notice evaluative differences as far as gender is concerned?

To try and answer these questions, section 2 briefly reviews the literature; section 3 introduces the methodological approach — corpus linguistics — and the tools involved in the selection of the data, before tackling the reasons for the choice of the English Web 2020 corpus (enTenTen 20, see SketchEngine 2022). Section 4 focuses on the semantic preference of a selection of nominal collocations used to refer to AP, and section 5 addresses the question of the generally perceived negative axiology linked to the discourses on old age and aging through the notion of semantic prosody. The potential gender differences are also examined through the study of the collocations [adjective + *old* / *older* / *elderly* / *middle-aged* / *mature* + noun], as developed in sections 4 and 5, following Mautner’s “corpus linguistic methods within a sociolinguistic framework” (Mautner 2007, 51) to establish collocational profiles for each collocation. In line with this author, we posit that “age is discursively constructed” (53, see also Hareven 1995) and that labelling expressions play a crucial role in the categorization of groups of people:

If we accept that labeling plays a crucial role in categorization, boundary drawing, and stereotyping, then the study of age labels emerges as a worthwhile contribution to aging research (Mautner 2007, 53).

While this study mainly focuses on old age and aging, we also aim to discuss potential evaluative gender differences. Gender is a complex and multifaceted notion that has been widely described and studied, especially in the field of sociolinguistics where it has been regarded as one of the factors for language variation and change, but also for perception differences. It should be noted that we deliberately use the term “gender” to refer to the expression of constructed social identity, which has to be distinguished from “sex” pointing only to biological identity (Cheshire 2004; Eckert 1989).

2 Literature review and research questions

The topical nature of the study has to be noted, as there is a growing interest in age-related questions due to growing populations of aging people all over the world. This

2. Semantic prosody is also referred to as “evaluative prosody,” “discourse prosody,” “emotive prosody,” “pragmatic prosody” or “attitudinal prosody.”

resulted in various and varied studies within CARE and/or medical contexts (Escourrou et al. 2022, Abdi et al. 2019, Fulmer et al. 2021, etc.), but strangely enough, in relatively few studies on the language used to talk about AP *per se*. Notable exceptions are Tommaso's (2015) work on media depictions of age and aging, Coupland, Coupland, and Giles's (1991) *Language, Society and the Elderly*, Coupland's (1997) study on language, aging and agism, Wangler and Jansky's (2023) media portrayal of old age, etc., and most specifically Mautner's (2007) work on collocational profile of the word "elderly." Following Gerstenberg and Lindholm's (2019, 1) call to an "interaction and collaboration between diverse disciplines [...] to allow for collaborations across these barriers," this paper will focus on the linguistic aspect of some sequences used to refer to AP, and the impact these denominations have on the representations of old age and aging in terms of evaluation (see Appraisal theory), but also regarding the potential gender differences. Following one of the main tenets of Corpus Linguistics, i.e. the Firthian conception of contextual meaning or Sinclair's principle of "no independent word meaning," we posit that the meaning of words in actual occurrences of language is rarely the dictionary meaning, but a discursive meaning not generated by the sum of the words following one another, but rather by the interaction of those words together (Firth 1957, Channel 2000, Sinclair 2004, etc.); this means that the meaning of a given item is "frequently shared across units in discourse" (Morley and Partington 2009, 139) or, as Firth (1957, 179) writes, "You shall know a word by the company it keeps," especially when it comes down to evaluative meaning. To put it simply, and metaphorically, "elements of meaning 'hunt in packs.'" (Morley and Partington 2009, 140). Semantic preference and semantic prosody are two concepts frequently resorted to in Corpus Linguistics, but they are also often confused.³ Semantic preference is defined by Stubbs (2001, 65) as "the relation, not between individual words, but between a lemma or word-form and a set of semantically related words." A given item shows semantic preference when it co-occurs with "a class of words which share some semantic feature" Stubbs (2001, 88). Semantic prosody, a.k.a. "evaluative prosody," "discourse prosody," "emotive prosody," "pragmatic prosody" or "attitudinal prosody," is a useful—though sometimes controversial and contentious (Whitsitt 2005, Hunston 2007, Bednarek 2008, Stewart 2010)—notion in Corpus Linguistics⁴ that was introduced by Louw,⁵ who defines it as "a consistent aura of meaning with which a form is imbued by its collocates" (Louw 1993, 157). As Soto-Almela and Alcaraz-Mármol (2017, 153) state, "In 2000, Louw amplifies his definition of semantic prosody going beyond this 'aura of meaning' (1993) to clearly distinguish it from connotation." Semantic prosody is thus not just synonymous with connotation, even if some scholars refer to it as "connotational meaning" (Morley and Partington 2009, 150) or integrate the term "connotation" within its definition (Bednarek 2008).⁶ Several studies have already been conducted

3. For further reading on the differences between semantic preference and semantic prosody, see Partington (2004), Bednarek (2008) and more recently Begagic (2018).

4. For a comprehensive analysis of the very concept of "semantic prosody," see Hunston (2007), Bednarek (2008).

5. More precisely, Louw (1993, 158) recognizes that the term has been coined by Sinclair during a personal communication (1988), who himself borrowed it from Firth (1957) who used the term "prosody."

6. "If we want to keep the terms *semantic prosody* and *semantic preference*, we should reserve the term *semantic preference* for collocations of lexical items with (more or less specific) semantic subsets, and use the term *semantic prosody* for connotations of all kinds. *Semantic prosody* then refers to POS/NEG connotation as well as more complex attitudinal connotations affecting both simple words

on semantic prosody, in various fields, such as politics, mostly immigration (Soto-Almela and Alcaraz-Mármol 2016, 2017; Jamet and Lafiandra 2023), mental health (Jamet 2018, 2022; Jamet and Coupé 2023), but, to the best of our knowledge, not on the language of old age and aging. This paper therefore intends to fill in a gap by focusing on the role played by semantic preference and semantic prosody in the contemporary denotation of AP in English, mostly through the adjectives used in combination within ten typical collocations, and to bring out the collocational profiles of expressions referring to AP. The hypothesis in this paper is that the meaning of a particular item or phrase cannot be brought out in isolation, but only within the discursive context—left and right—of the phrase, which creates a specific semantic preference for said item or phrase. Yet, it would be misleading to just consider the surrounding context to evaluate the semantic prosody of a given token, as the node also plays a role in the evaluation, as Morley and Partington (2009, 142) clearly state: “The relationship between the node and the surrounding items, as we said, needs to be contemplated in the definition.” Indeed, a clearly negatively or positively connoted node will affect the semantic prosody of the token under consideration, regardless of the influence of the surrounding context. That is why, depending on researchers, there are two schools of thought (Ben Ghazlen 2022): semantic prosody is either in the lexical item itself (items have, in Hoey (2005)’s terminology, lexical primings, i.e., a set of suggestions on how to use them, on how they normally interact with other items), or in the overall discourse (Hunston 2007). We will depart from the two anti-nomic positions, following Morley and Partington (2009, 156), and consider that “item and environment do interact and affect each other, that repeated usage of an item in new environments will alter the priming instructions-suggestions of the item itself,” and that “some words form an evaluative meaning ‘pack’” (140).

3 Methods and data

This section discusses the corpus-based method implemented for the study of the evaluative representations of AP. Section 3.1 provides some data on the corpus and section 3.2 expands on the analyses carried out on the corpus.

3.1 Corpus selection

As our intention is not to examine the evaluative evolution of our items in a diachronic perspective, we adopted a purely synchronic point of view, which means that we are only interested in the contemporary state of the language. This theoretical decision is reflected in the methodology and choice of the corpus to collect our data: we opted for the English Web 2020 corpus, a.k.a. enTenTen20, a reference corpus available on SketchEngine (Kilgarriff et al. 2014; SketchEngine 2022). The corpus contains approximately 43 billion tokens (36 billion words); see Appendix Figures 5 and 6). The English Web 2020 has been selected for this study because it is considered as a reference corpus and, as such, it is balanced and representative of contemporary English (2019–2021). It is composed of a variety of sub-corpora from different English-

and larger units of meaning such as phrases, i.e. concerns both ‘traditional’ connotation (said to relate to single words) and ‘prosodic’ connotation (connotation that is “distributed prosodically across a textual sequence” (Stubbs 2001, 202).

speaking countries (US, UK, Canada, Australia, NZ, etc.), and contains different registers (arts, business, games, health, home, recreation, reference, science, etc.; see Appendix Figures 6 and 7). Mautner also used a large corpus, the 500-million-word *Bank of English*. *Wordbanks Online*, and justified her choice as follows:

[L]arge reference corpora are rich repositories of social information and thus offer considerable potential for research in sociolinguistics and discourse analysis. [...] Millions of words from diverse genres allow one to be reasonably confident about generalizability, especially when the item under investigation, as it is the case with *elderly*, has a high total frequency in the corpus concerned. (Mautner 2007, 52, 65)

3.2 Analyses carried out

To evaluate semantic preference and semantic prosody, we analyzed the collocations containing a near-synonymous adjective related to old age followed by a noun referring to a human being. Xiao and McEnery (2006, 108) define synonyms and near synonyms as follows: “[lexical pairs] that have very similar cognitive or denotational meanings, but which may differ in collocational or prosodic behavior. As such, synonymous words are not collocationally interchangeable.” Hu confirms this non-interchangeability:

In particular, even though two words may share similar cognitive or denotational meanings, they may demonstrate not only different collocational behavior but also distinct semantic prosodies. (Hu 2015, 118)

To do so, we adopted a statistical method, conceptualized from a discourse standpoint, which entails “analysing, via a concordance, how a node is actually instantiated many times in many texts” (Partington 2015, 292). As reminded by Ben Ghazlen:

Specifically, by exploring the patterns of co-occurrence that are evidently inaccessible to the naked eye, the researcher can statistically pin down the type of semantic polarity that prevails and the evaluative force that the node acquires as a result of its collocational behavior. In biogenetics terms, this prosodic behavior will be instantiated in the DNA of the item, and progressively built up in the minds of speakers [...] by the number of exposures to uses. (Ben Ghazlen 2022, 64)

To retrieve the most frequently used sequences of words in the English Web 2020 corpus, we used Complex Query Language (CQL), “a special code or query language used in Sketch Engine to search for complex grammatical or lexical patterns or to use search criteria which cannot be set using the standard user interface” (SketchEngine 2017–2024). The CQL searches were carried out using the attributes *tag* for part of speech, with the value “J.*” representing a sequence of adjectives, *word* for word-forms, and *lemma* for lemmatized forms.⁷ Additionally, the question mark wild card ? for one unspecified character, and the asterisk * for any number of a specified character were used.⁸ A random selection of 10 000 occurrences for each query was

7. For example, sequences of adjectives followed by specific nouns: [tag="J.*"] [word="people"], or [tag="J.*"] *man* / *men*, or [tag="J.*"] [lemma="man"], or [tag="J.*"] [lemma="woman"].

8. For example, *m?n* finds both *man* and *men*, and *wom?n* finds both *woman* and *women*.

carefully examined, and the most frequent sequences used to refer to AP were retrieved (Table 1).

Table 1: Adjectives used to refer to AP in enTenTen 2020 with absolute and normalized frequency per million tokens (pmt)

| PEOPLE | | MAN / MEN | | WOMAN / WOMEN | | |
|---------------------------------|--------|--------------------------------|-----------|-------------------------------|---------------|-----|
| 11 203 058 tokens 259.78 pmt | | 5 686 569 tokens 131.86 pmt | | 3 693 404 tokens 85.64 pmt | | |
| aged | Ø | | man | 3 | woman | 2 |
| elder | Ø | | man / men | 6 | woman | 3 |
| elderly | people | 7 | man / men | 60 | woman / women | 89 |
| gray- / grey-haired | Ø | | man | 5 | woman | 4 |
| gray-headed | Ø | | man | 1 | Ø | |
| mature | people | 3 | man / men | 5 | woman / women | 13 |
| middle-aged | Ø | | man / men | 22 | woman / women | 32 |
| old | people | 22 | man / men | 660 | woman / women | 232 |
| older | people | 43 | man / men | 90 | woman / women | 181 |
| oldest | Ø | | man / men | 6 | women | 2 |
| senior | Ø | | men | 10 | woman / women | 21 |
| white-haired | Ø | | man / men | 3 | woman | 2 |
| | | 75 | | 871 | | 581 |

The adjectives shared by *people*, *man* and *woman* are respectively *old*, *older*, *elderly* and *mature*. Given their high frequency, we decided to focus more specifically on those four adjectives *old* / *older* / *elderly* / *mature* followed by what will be termed the head noun, i.e., a noun referring to a human being or a group of human beings: *man* / *men* or *woman* / *women* or *people*. We nevertheless added to the list the adjective *middle-aged* which seemed quite productive with *man* and *woman*, even if it was not found in collocations with *people* in the first 10 000 rows.⁹ This selection of five adjectives resulted in 25 combinations.¹⁰ To retrieve these sequences, we performed a basic search in the English Web 2020 corpus. Figure 1 is a visual representation of the results of the most frequent lemmatized forms (see Appendix Table 4; for the most frequent non-lemmatized forms, see Table 3).

As shown in Figure 1, depending on the head noun, the most frequently used adjective is either *old* or *older*, which are clearly the two most frequent adjectives in terms of normalized frequency, followed by *elderly*, *middle-aged* and *mature*, the last two showing a relatively low productivity when combined with the head nouns *m?n*, *wom?n* and *people*. In terms of frequency, the most frequently used phrases containing an attributive adjective are [adjective + old m?n], [adjective + old wom?n], [adjective + older m?n], [adjective + older wom?n], showing a clear predominance of *old* and *older* used to premodify *m?n* and *wom?n*. *Middle-aged* and *elderly* show a relatively lower frequency of use, either in combination with *m?n* and *wom?n*. *Mature* is relatively infrequent compared to the other adjectives under scrutiny. The second

9. We did not include the adjective *senior* in our analyses as it was not tagged as an adjective in SketchEngine but as a noun. Further research may focus on the collocational profile of expressions including *senior*.

10. *old woman*, *old women*, *old man*, *old men*, *old people*; *older woman*, *older women*, *older man*, *older men*, *older people*; *elderly woman*, *elderly women*, *elderly man*, *elderly men*, *elderly people*; *mature woman*, *mature women*, *mature man*, *mature men*, *mature people*; *middle-aged woman*, *middle-aged women*, *middle-aged man*, *middle-aged men* and *middle-aged people*.

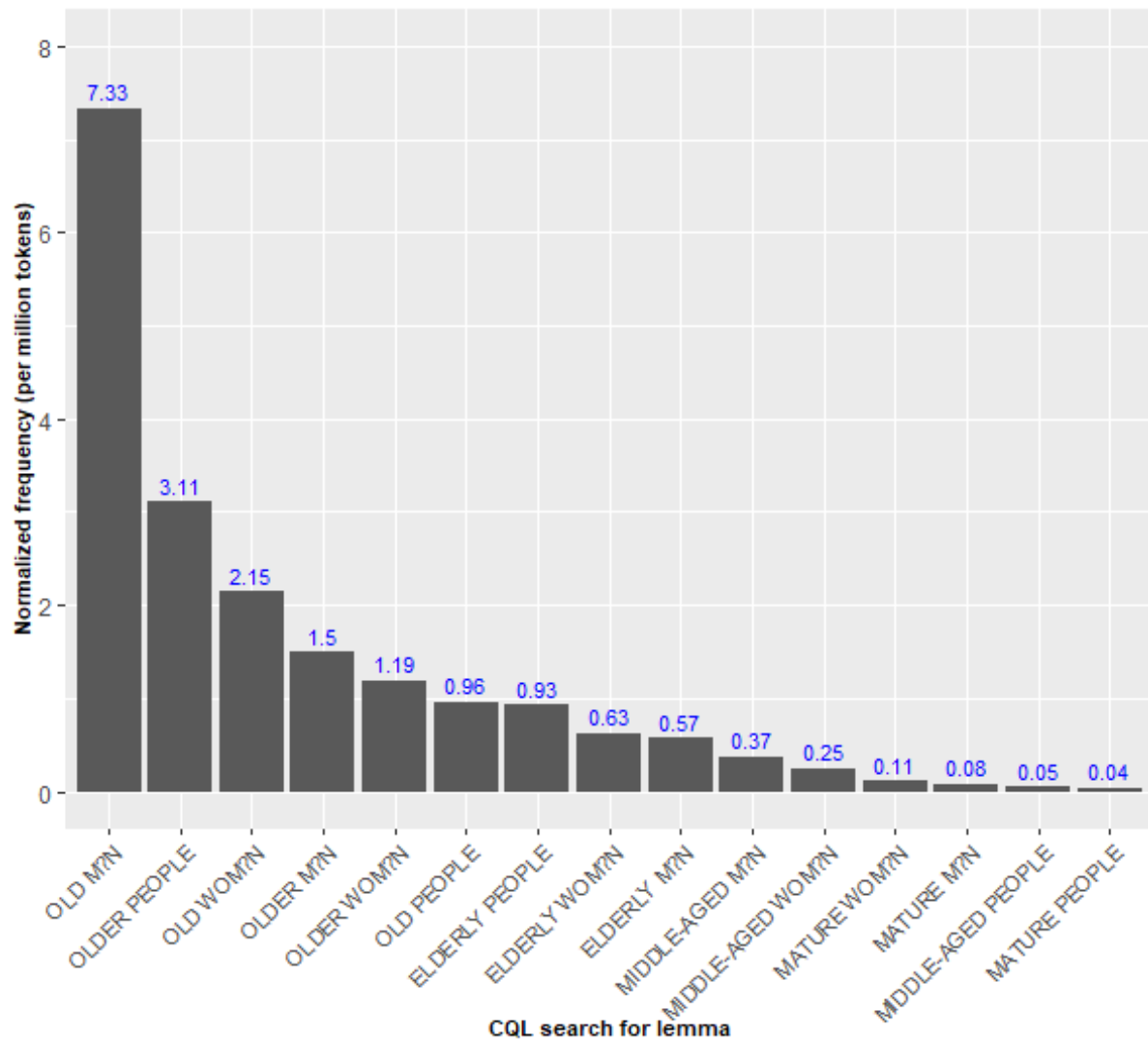


Figure 1: Normalized frequencies of sequences to refer to AP

type of CQL searches was intended to retrieve the most frequently used adjectives premodifying the sequences under scrutiny, so as to analyze the axiology conveyed by the attributive adjectives. To retrieve the sequences [adjective + *old* / *older* / *elderly* / *middle-aged* / *mature* + head noun] and analyze in a gender perspective the evaluation through the adjective used, we performed several CQL searches.¹¹ Looking at the adjectives preceding the selected phrases was deemed an appropriate way of studying their semantic preference and semantic prosody through the choice of adjectives, as reminded by Ben Ghazlen (2022, 68), “[u]ndeniably, the study of the semantic prosody of a given term basically implies the extraction and scrutiny of its typical collocates.” The results are exemplified in Appendix Table 5, showing the CQL searches carried out and ranked by normalized frequencies.¹² The first type of analyses carried out was to consider the most frequent adjectives used in combination with the head nouns.

4 Corpus investigation: semantic preference of words referring to AP

This section presents the results of the quantitative and qualitative corpus-based analyses carried out through AntConc (Anthony 2019). We will consecutively analyze and comment on each sequence of adjective + *old* / *older* / *elderly* / *middle-aged* / *mature* + *m?n* / *wom?n* to see whether similarities and/or differences can be pointed out regarding semantic preference.

4.1 *old m?n and old wom?n*

The Free USAS English web tagger (Rayson 2023) was used to analyze the semantic preference for the sequences under scrutiny. This tool was chosen as it enables researchers to tag semantically each adjective used in combination with said sequences. Some semantic tags (from now on, “Semtags”) were considered erroneous, for two main reasons: (1) the word was not recognized by USAS (Rayson 2023), and tagged Z99 “Unknown,” such as *crotchety*, *yr* (for *year*), *wizened*, etc.; to be able to label unknown words with the appropriate Semtag, we used various synonyms, and when the tagger labeled the word with the same Semtag, we used this Semtag; (2) due to polysemy, some words were inappropriately tagged.¹³ The most frequent semantic domains (Semtags) for the adjectives used in attributive position to premodify *old m?n*

11. For example, the query [tag="J.*"] [word="old"] [lemma="man"] retrieves all occurrences of sequences containing an adjective followed by the word *old* and the lemma MAN, finding occurrences of *man* and *men*.

12. To analyze the most frequently used adjectives, the Excel file generated with the CQL searches was cleaned, and some occurrences were removed for two main reasons. The first reason was due to tagging mistakes by SketchEngine (2014): words such as *many*, *several*, *few*, *a few*, *fewer*, *such* were tagged as adjectives, and not as determiners (predeterminer for *such*) by SketchEngine (2014); *only* and *very* were systematically tagged as adjectives, even when they were used as adverbs; *much* and *less* were tagged as adjectives, when they were used as adverbs premodifying the adjective, etc. The second reason was due to typing mistakes, such as *lobeAn*, i.e., most frequently words stuck together and tagged as adjectives because of their prenominal position. Once the Excel documents were cleaned, other files were created to keep only the adjectives modifying the expressions referring to AP.

13. For example, *sweet*, which was tagged X3.1 “Sensory: Taste,” and which was changed to O4.2+ “Judgment of appearance;” *straight*, which was tagged O4.4 “Shape” and *hot*, which was tagged O4.6+ “Temperature,” and which were both changed to S3.2 “Relationship: intimate / sexual;” *single*, which was tagged N5 “Quantities,” and which was changed to S4 “Kin,” etc., see italics in Appendix Table 11.

(see Appendix Table 6) are respectively Emotional actions, states and processes (E), Substances, materials, objects and equipment (O), and Social actions, states and processes (S). This is in keeping with the most frequently used attributive adjectives used in collocation with the sequence *old m?n* (see Appendix Table 11 (a) and Figure 8 (a)). They are essentially related to mental, psychological characteristics (*grumpy, wise, grand, cranky, bitter, lonely, strange, creepy, crotch, vulnerable, nice*, etc.). The physical characteristics are rarely alluded to with *old m?n*, except with *bearded* and *fat*, which exhibit a relatively low frequency of use (to which could be added *dirty*, referring both to physical and psychological properties), as illustrated in (1) and (2):

- (1) These are his adventures in that innocent time long before he became the **grumpy old man** with white hair and a stick. (pearsecom.co.uk)
- (2) The archetypes that Gandalf epitomizes include those of the wizard or magician, the **wise old man** who acts as a teacher and/or manipulator and whose principle role is to see that the major actors complete their quest. (silmarillionwriters-guild.org)

The most frequent semantic domains (Semtags) for the adjectives used in attributive position to premodify *old wom?n* (see Appendix Table 11 (b) and Appendix Figure 8 (b)) are somehow different from those found with *old m?n*: General and abstract terms (A), Substances, materials, objects and equipment (O), and Social actions, states and processes (S). As for *old m?n*, the adjectives used in attributive positions are mostly linked to mental, psychological characteristics (*poor, wise, lonely, crazy, bitter, strange, frail, good, mysterious, grumpy, silly, eccentric, wicked, rich*, etc.), but there seems to be a tendency to also refer to physical characteristics for *old wom?n* (*ugly, fat, toothless, wizened* (3)), even if the majority of occurrences refer to mental, psychological attributes ((4), (5) and Appendix Table 11).

- (3) “Come here, Aunt Milly,” she called out, and she sat down on the highest step and waited till the **fat old woman**, enveloped in a coarse gray blanket, joined her. (gutenberg.org)
- (4) A large group of college students gathered around a **wise old woman** as she spoke wisdom into their lives. Now in her eighties, she had a wealth of knowledge and experience from a life well. (christianitytoday.com)
- (5) He had heard people speak of her as “a **wicked old woman**.” Perhaps she was inside the carriage... but he only saw the Castle coachman and footman and the coronet on the door. (readcentral.com)

The semantic preference for adjectives used attributively with *old m?n* and *old wom?n* is thus relatively similar, showing no real gender differences for the expressions containing *old*. We will now see if there are any differences for *older m?n / wom?n*.

4.2 Older m?n and older wom?n

The most frequent semantic domains (Semtags) for the adjectives used in attributive position to premodify *older m?n* (see Appendix Table 7 (a)) are Social actions, states and processes (S), Substances, materials, objects and equipment (O), and General and abstract terms (A). The characteristics conveyed by the adjectives used to modify *older m?n* (see Appendix Table 12 (a) and Appendix Figure 9 (a)) are more positive than those used with *old m?n*: *healthy*, *wealthy*, *nice*, etc. They refer both to positive mental, psychological characteristics (*powerful*, *mysterious*, *wise*, *lonely*, *nice*, *distinguished*, etc., (6)) and physical characteristics (*handsome*, *attractive*, *white-haired*, *sexy*, *bearded*, *tall*, *hot*, *charming*, etc., (7)). The adjectives insist on stability, be it financial / economic (*wealthy*, *rich* (8)), intellectual (*powerful*, *distinguished*, *successful*, *charismatic*, *experienced* (9)) or marital (*married* (10)), in sharp contrast with *elderly*. This positive aspect is reflected in the main semantic domains mentioned above.

- (6) At that moment Jung envisioned a **powerful older man** within who he described as an elderly 18th Century gentleman of wealth and position, sporting a white wig. (jungcircle.com)
- (7) feeling naughty (Quincy, 02171, Norfolk County) feeling naughty and looking for a **sexy older man** to satisfy my urges, I will come bi curious females you and be your toy. (curiousbi.com)
- (8) [...] numerous students have joined the Sugar Daddy “dating” website, where **wealthy older men** meet young women, or “Sugar Babies,” helping them with university fees and rent. (stuff.co.nz)
- (9) George Gallo (Director/Writer) uses the voice of an **experienced older man**, who is a master artist, to express his feelings to a young artist about his views of life and art, which equals the wisdom [...] (bestmoviereviews.org)
- (10) “She wants the money/ It comes with his cream,” Steen sings, spinning the tale of a gold-digging young woman who sleeps with a **married older man** in exchange for money, handbags, and other largess. (tinymixtapes.com)

The most frequent semantic domains for the adjectives used attributively with *older wom?n* (see Appendix Table 7 (b)) are Substances, materials, objects and equipment (O), Social actions, states and processes (S), and General and abstract terms (A), in other words the same ones used with *older m?n*, though not in the same order for the first two. The characteristics conveyed by the adjectives used to modify *older wom?n* (see Appendix Table 12 (b) and Figure 9 (b)) are similar to those used with *older m?n*: they are more positive than those used with *old wom?n*: *healthy*, *wealthy*, *attractive*, *rich*, etc., (11). They refer both to positive mental, psychological characteristics (*mysterious*, *nice*, *sweet*, *lonely*, *wise*, *lovely*, *experienced*, *powerful*, etc., (12)) and physical characteristics (*attractive*, *beautiful*, etc., (13)). The adjectives also insist on stability,

be it financial / economic (*wealthy, rich* (14)), intellectual¹⁴ (*experienced* (15)), but not marital (*single* (16)). This predominantly positive aspect is reflected in the main semantic domains mentioned above.

- (11) [...] Kyle goes outside to check his mailbox. On his way to the mailbox he spots an **attractive older woman** stepping outside her apartment in a white robe. Kyle smiles at the woman. (sdc.com)
- (12) [...] Deborah the future judge of Israel becomes the prototype of the questing heroine. Taught by a **mysterious older woman** to develop spiritual power by meditating on the text “How good it is for people to live in peace,” Deborah is both seer and leader (awpwriter.org)
- (13) Ivar the Boneless had suddenly lost his lust for battle, and was now just staring blankly at the **beautiful older woman** in front of him. Bjorn Ironside glanced down at his mother, Lagertha, who stood still in thought. (roleplaygateway.com)
- (14) In Agatha Christie’s theatrical masterpiece, Leonard Vole is arrested for the murder of Miss Emily French, a **wealthy older woman**. Unaware that he was a married man, Miss French made him her principal heir, casting suspicion on Leonard. (theatreworks.us)
- (15) [...] an invitation for Dom training could be exactly what he needs to regain his self-confidence. Challenged with an **experienced older woman**, he soon realizes wielding a flogger isn’t always enough to break down the emotional walls [...] (nightowlreviews.com)
- (16) Reveals that almost one-quarter (24%) of Canadian women are raising children on their own and 14% of **single older women** are poor. Describes barriers and opportunities for young women. (cwhn.ca)

The semantic preference for adjectives used attributively with *older m?n* and *older wom?n* is thus relatively similar, showing no major gender differences, and relatively positive; this can be accounted for by the fact that *older people* are not necessarily *old*. We will now see that there are differences for *elderly m?n / wom?n*.

14. To a lesser extent than for *older m?n* yet.

4.3 *Elderly m?n and elderly wom?n*

The most frequent semantic domains for the adjectives used to premodify *elderly m?n* (see Appendix Table 8) are: General and abstract terms (A), The body and the individual (B), and Social actions, states and processes (S). The adjectives used (see Appendix Table 13) seem more varied and tend to be used in more specialized contexts: *missing*, *unidentified*, *disabled*, *local*, etc., (17) and (18)). They mostly refer to health (*frail*, *vulnerable*, *disabled*, *sick* (19)), i.e. to a position of vulnerability, as developed by Mautner (2007):

- (17) This call out was initiated by Greater Manchester Police. 10TH Oct: A full scale search was conducted for a **missing elderly man** in the Mirfield area. Support was provided by Calder Valley SRT and also a Trailing dog was requested from Cheshire SAR. (holmevalleymrt.org.uk)
- (18) Salt Lake County Health Department spokesman Pam Davenport confirmed the gender Monday of first person in the continental United States to die while infected with the virus. The **unidentified elderly man** suffered symptoms of the disease and died after traveling abroad to an area with a Zika outbreak. (medicalxpress.com)
- (19) A meeting was arranged for the following Wednesday at the sick man's home. We went upstairs and found this very **sick elderly man** lying in bed with boxes of medication lying around him and a large oxygen bottle nearby. (southallchristianfellowship.org.uk)

The tendencies for *elderly wom?n* are very similar to those found with *elderly m?n*: the most frequent semantic domains for the adjectives used to premodify *elderly wom?n* (see Appendix Table 8) are identical: General and abstract terms (A), The body and the individual (B), and Social actions, states and processes (S). The adjectives used (see Appendix Table 13) also seem more varied and tend to be used in more specialized contexts, and are the same as the ones found with *elderly m?n*: *missing*, *unidentified*, *disabled*, *local*, etc., (20). The same adjectives used with *m?n* and *wom?n* mostly refer to health (*frail*, *vulnerable*, *disabled*, *sick*, *confused*), and more particularly to the position of vulnerability previously mentioned (21):

- (20) STARICA N., an **unidentified elderly woman**, slaughtered in Suva Reka. 260. STARICA N., another **unidentified elderly woman** slaughtered in Suva Reka 261. TALIC MIRKO (age 70) and 262. TALIC JOVANKA (age 65), from Strpce, were massacred [...] (kosovo.net)
- (21) Gardaí are hunting fraudsters who are believed to have driven a **confused elderly woman** with dementia to the bank to withdraw almost 1,000 in recent weeks. The frail 87-year old mistakenly drew 90 initially out of her Bank of Ireland account but was immediately sent back to the cashier by the con artists to withdraw ten times that amount. (breakingnews.ie)

The semantic preference for *elderly m?n* and *elderly wom?n* is very similar and the connotation was initially quite neutral, probably due to the euphemistic origin of *elderly* to refer to AP, even if the passivity and vulnerability of *elderly people* is generally highlighted. This is in sharp contrast with the semantic preference of *older m?n* and *older wom?n*. The analyses confirm Mautner's results regarding the semantic preference of *elderly*: "Dominant semantic preferences include disability, illness, care, and vulnerability to crime, with the associated semantic prosodies often negative" (2007, 63). Yet, the connotations we found for *elderly* are not so negative in our corpus, which reminds us of Nuessel's hesitation between the neutral, non-stereotyping qualities of *elderly* (Nuessel 1982) and the negative connotation he called "stigmatic because this word has been employed by media reporters who have traditionally portrayed this social subdivision in a negative and derisive fashion" (Nuessel 1984–1985, 18). Let us now consider the semantic preference for *middle-aged m?n / wom?n*.

4.4 *Middle-aged m?n and middle-aged wom?n*

The most frequently used semantic domains for the attributive adjectives modifying *middle-aged m?n* (see Appendix Table 9) are: General and abstract terms (A), Substances, materials, objects and equipment (O), and Numbers and measurement (N). The adjectives (see Appendix Table 14) mostly refer to physical characteristics (*white, looking, balding, handsome, dressed, tall, etc.*, (22) and (23), with an emphasis on weight (*overweight, portly, fat, obese, stout, (24)*), a fact confirmed by the frequency of the N semantic domain:

- (22) He doesn't really have any remarkable features, physically speaking that is. A grey, **balding middle-aged man** with glasses and with an alleged dark side is the best way to describe him. (smh.com.au)

- (23) Madeleine accompanies me to the Youth Employment Bureau at the Town Hall in Crouch End, run by a Mr. Kingswell, a grey-haired **handsome middle-aged man** of above average height. (militantesthetix.co.uk)

- (24) Kymon is a tall, slightly **overweight middle-aged man** with sparse grey hair, intelligent, dark eyes, a short-trimmed beard and a beautiful, warm, baritone voice. (pandius.com)

- (25) Towards the beginning of the story, the reader gets the impression that she is an **ordinary middle-aged woman** who's living a normal life. However, this changes as the main character learns about the case of a missing girl. (termpaperwarehouse.com)

- (26) There was a pause. Maggie, a little calmer, realised Grace, who had sunk into a chair. She saw that **stout middle-aged woman** with the flat expressionless face and the dull eyes. She saw the flabby hands nervously trembling (readcentral.com)

- (27) Also, everyone swooning over Palin needs to get out more. She's a reasonably **attractive middle-aged woman**, which makes her way hotter than every other Republican in national office, but sheesh. (unfogged.com)
- (28) Elizabeth Douglas might have appeared to Ann as being a somewhat overweight and **plain middle-aged woman**, indistinguishable from so many such women who frequented the loyalist bars at the weekend. (douglashistory.co.uk)

The semantic preference for *middle-aged m?n* and *middle-aged wom?n* is similar on the whole, even if it is more negative for *wom?n* than for *m?n*. We will now see that the trend is reinforced for the last sequence, *mature m?n* / *wom?n*.

4.5 *mature m?n and mature wom?n*

The most frequently used semantic domains for the adjectives premodifying *mature m?n* (see Appendix Table 10) are: Social actions, states and processes (S), General and abstract terms (A), and Emotional actions, states and processes (E). The adjectives (see Appendix Table 15) are rather positive, and refer to mental characteristics (*serious, strong, qualified, confident*, (29)) but also to physical characteristics, with sexual connotations (*active, horny, sexy*, (30)):

- (29) That's why i am looking for someone who will give me wings to fly =) I am looking for a **serious mature man**, a real defender for me, a man who knows what he wants in life. I am looking for someone that still knows how to trust people (step2love.com)
- (30) Many fathers are not reliable and you will be sure right after watching these movies where **horny mature men** are penetrating their sons girlfriends. No incest and only exclusive sons girlfriend fucking scenes. (smut-blogs.com)

The most frequent semantic domains for the adjectives used in combination with *mature wom?n* (see Appendix 26) are somehow different from those used with *mature m?n*: Substances, materials, objects and equipment (O), Social actions, states and processes (S), and General and abstract terms (A). Depending on the point of view, the adjectives (see Appendix Table 15) can be seen as positive, or extremely negative, as they tend to portray women as sexual objects (*sexy, hot, attractive, nude, busty, chubby, wild*, (31)), with a strong emphasis on their physical attributes (*beautiful, attractive, gorgeous*, (32)):

- (31) Not Enough Cheating Wives Caught on Tape. I love **nude mature women**: Pics, vids cams — any type of housewife porn a site can offer. (milfsexreports.com)
- (32) love to join you and hubby for some joy. where are you from?? e-mail me if you would like some pics. Beautiful sexy drop dead **gorgeous mature woman** with tits that are truly awesome!! (xelyd.com)

Therefore, there appears to be a significant gender difference for *mature m?n* and *mature wom?n* as exemplified by the adjectives used in combination: if the adjectives used with *mature m?n* are positive and refer both to mental and physical attributes, those used with *mature wom?n* are rather negative and mostly insist on physical properties, especially sexual ones. This gender difference leads us to finally consider the role of semantic prosody in the representations of aging men and women, and to assess its effect in relation to the semantic preference of said collocations.

5 Corpus investigation: semantic prosody of words referring to AP

In this final section, special emphasis is laid on the role of semantic prosody in the representations of AP. The previous analyses showed that there are no real gender differences for *elderly* and *old*, which is the contrary for *middle-aged*, *mature* and *older*. One reason for that could be the fact that *the old* and *the elderly* can be used as nominalized adjectives referring to an entire, indistinct group of AP, which is not the case for **the mature*, **the middle-aged* or **the older*. The final analysis consisted in aggregating the total number of adjectives used in combination with the headwords, to see whether some general tendencies could be pointed out. Table 2 (a) and Table 2 (b), and Figures 2 (a) and 2 (b) show the results for *m?n* and *wom?n*.



Figure 2: Word cloud of most frequent adjectives used in attributive position with *old / older / elderly / middle-aged / mature* (a) *m?n* and (b) *wom?n*

The adjectives used to premodify men of a certain age mostly insist on mental characteristics: *grumpy*, *wise*, *grand*, *good*, *crazy*, *cranky*, *bitter*, *lonely*, *strange*, *creepy*, *angry*, *crotchety*, *nice*, *venerable*. Yet, those characteristics seem predominantly negative, with 13 negatively connoted adjectives, 5 positively connoted adjectives and 3 neutral adjectives, contrary to our initial intuition.

The adjectives used to refer to aging women also insist on mental characteristics (*wise*, *lonely*, *wealthy*, *mysterious*, *rich*, *sweet*, *crazy*, *strange*, *bitter*, *eccentric*, *good*) but also on physical attributes (*frail*, *beautiful*, *attractive*, *ugly*, *healthy*, *looking*). The adjectives are surprisingly more positive than for men: 8 positively connoted adjectives, 8 negatively connoted adjectives and 3 neutral adjectives.

An analysis of the most frequent semantic domains of the adjectives used in attributive position with Wmatrix5 (Rayson 2009) confirms the tendency to find both

negative and positive evaluations for *m?n* and *wom?n*, as illustrated in Figures 3(a) and 3(b), which show the evaluative semantic domains shared by both: Evaluation: _Good; Evaluation: _Bad; Judgment_of_appearance: _Positive; Judgment_of_appearance: _Negative.

Table 2: Most frequent adjectives used in attributive position with (a) *old/older/elderly/middle-aged/mature m?n* and (b) *old/older/elderly/middle-aged/mature wom?n*

| Type | NormFreq | Rank | Freq | Semtag | Semantic prosody | Semantic domain |
|---------|-----------|------|------|----------|------------------|---------------------------------------|
| year | 67162.737 | 1 | 695 | T1.3 | Neutral | Time: Period |
| dirty | 47158.871 | 2 | 488 | O4.2- | Negative | Judgement of appearance (pretty etc.) |
| grumpy | 38364.901 | 3 | 397 | E3- | Negative | Calm / Violent / Angry |
| little | 38074.990 | 4 | 394 | A13.7 | Neutral | Degree: Minimizers |
| wise | 31793.583 | 5 | 329 | S1.2.6+ | Positive | Sensible |
| grand | 31696.946 | 6 | 328 | O4.2+ | Positive | Judgement of appearance (pretty etc.) |
| poor | 25705.450 | 7 | 266 | I1.1- | Negative | Money: Affluence |
| good | 18361.036 | 8 | 190 | A5.1+ | Positive | Evaluation: Good / bad |
| crazy | 13239.273 | 9 | 137 | B2- / X1 | Negative | Health and disease / General |
| cranky | 11886.355 | 10 | 123 | A6.2- | Negative | Comparing: Usual / unusual |
| bitter | 11306.533 | 11 | 117 | O4.2- | Negative | Judgement of appearance (pretty etc.) |
| lonely | 10630.073 | 12 | 110 | S5- | Negative | Groups and affiliation |
| frail | 10146.888 | 13 | 105 | S1.2.5- | Negative | Toughness; strong / weak |
| strange | 8697.333 | 14 | 90 | A6.2- | Negative | Comparing: Usual / unusual |
| creepy | 8504.059 | 15 | 88 | O4.2- | Negative | Judgement of appearance (pretty etc.) |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic prosody | Semantic domains |
|------------|------------|------|------|---------|------------------|---------------------------------------|
| year | 117365.752 | 1 | 2037 | T1.3 | Neutral | Time: Period |
| little | 37335.792 | 2 | 648 | A13.7 | Neutral | Degree: Minimizers |
| poor | 33302.604 | 3 | 578 | I1.1- | Negative | Money: Affluence |
| other | 15268.495 | 4 | 265 | A6.1- | Neutral | Comparing: Similar/different |
| wise | 13770.454 | 5 | 239 | S1.2.6+ | Positive | Sensible |
| lonely | 12733.349 | 6 | 221 | S5- | Negative | Groups and affiliation |
| frail | 11408.159 | 7 | 198 | S1.2.5- | Negative | Toughness; strong / weak |
| beautiful | 10831.989 | 8 | 188 | O4.2+ | Positive | Judgement of appearance (pretty etc.) |
| attractive | 10371.053 | 9 | 180 | O4.2+ | Positive | Judgement of appearance (pretty etc.) |
| ugly | 8815.395 | 10 | 153 | O4.2- | Negative | Judgement of appearance (pretty etc.) |
| healthy | 8642.544 | 11 | 150 | B2+ | Positive | Health and disease |
| wealthy | 8354.460 | 12 | 145 | I1.1+ | Positive | Money: Affluence |
| single | 8066.375 | 13 | 140 | S4 | Neutral | Kin |
| mysterious | 7432.588 | 14 | 129 | A6.2- | Neutral | Comparing: Usual/unusual |
| looking | 7029.269 | 15 | 122 | A8 | Neutral | Seem / Appear |
| rich | 7029.269 | 15 | 122 | I1.1+ | Positive | Money: Affluence |
| sweet | 7029.269 | 15 | 122 | O4.2+ | Positive | Judgement of appearance (pretty etc.) |

(b)



Figure 3: Key domain cloud (Wmatrix, Rayson 2009) of most frequent adjectives used in attributive position with *old* / *older* / *elderly* / *middle-aged* / *mature* (a) *m?n* and (b) *wom?n*

Figure 4: Semantic prosody of adjectives used in attributive position with *old* / *older* / *elderly* / *middle-aged* / *mature* (a) *m?n* and (b) *wom?n*

Yet, as Figures 4(a) and 4(b) show, in terms of normalized frequency, the attributive adjectives used to premodify nouns referring to aging men are more negative (50 %) than those used to premodify nouns referring to aging women (26 %). If the positive semantic prosody is almost identical for men (23 %) and women (20 %), the main difference lies in the neutral semantic prosody (or absence of semantic prosody), which is unexpectedly much higher for women (54 %) than for men (27 %).

6 Concluding remarks and future investigations

This study attempted to show the relevance of combining the two concepts of “semantic preference” and “semantic prosody” to study the lexicalized sequences referring to AP in English and to bring out their collocational profiles. It brought to light the fact that apparently near-synonymous lexicalized sequences—*old* / *older* / *elderly* / *middle-aged* / *mature m?n* / *wom?n*—are not used interchangeably, and convey different usages, as exemplified in their different semantic preferences, as summarized by Turner (1998, 57): “Meaning is not a deposit in a concept container. It is alive and active, dynamic and distributed, constructed for local purposes of knowing and acting.” *Old* lays emphasis essentially on mental, psychological characteristics, when *older* can be used both for mental and physical characteristics, often in a positive way, with the notion of stability attached to it. *Elderly* is used in more specialized contexts, to insist on the passivity and vulnerability of AP, mostly in physical, health-related contexts. *Middle-aged* mostly focuses on physical attributes, more specifically weight, and is more negative when applied to women than men. Finally, *mature* is predominantly used for mental or physical characteristics, positively for men, but negatively for women because of the overtly sexual overtones. The gender differences depend on the sequences, with no real differences for *old* and *elderly*, minor differences for *older* and *middle-aged*, and significant differences for *mature*. Those gender differences apply to various semantic areas: physical vs. mental; sexualization vs. non-sexualization; positive vs. negative connotations, etc. Contrary to our initial intuition, sequences with *m?n* exhibit a more negative semantic prosody than those with *wom?n*, which are predominantly neutral. The same study needs to be carried out on the collective nouns referring to AP (*the old*, *the elderly*, *senior citizens*, etc.) to see if the same tendencies can be brought out, and whether the axiology is identical, regardless of any gender bias. A contrastive analysis of the different Englishes may also prove relevant to see if AP are represented similarly or differently in the different regions where English is spoken.

Ethics statement

We used the English Web 2020 corpus available on SketchEngine, and no personal data was collected at all. We referenced all quotations and provided the references in the bibliography, including the DOI when relevant.

Conflict of interest

The authors have no conflict of interest to declare.

References

- Abdi, Sarah, Alice Spann, Jacinta Borilovic, Luc de Witte, and Mark Hawley. 2019. "Understanding the care and support needs of older people: A scoping review and categorisation using the WHO international classification of functioning, disability and health framework (ICF)." *BMC Geriatrics* 19 (195). <https://doi.org/10.1186/s12877-019-1189-9>.
- Anthony, Laurence. 2019. *AntConc (Version 4.2.0) [Computer Software]*. <https://www.laurenceanthony.net/software>.
- Bednarek, Monika. 2008. "Semantic preference and semantic prosody re-examined." *Corpus Linguistics and Linguistic Theory* 4 (2): 119–134. <https://doi.org/10.1515/CLLT.2008.006>.
- Begagic, Mirna. 2018. "Semantic preference and semantic prosody—A theoretical overview." *Journal of Education and Humanities* 1 (2): 65–88. <https://doi.org/10.14706/jeh2018121>.
- Ben Ghazlen, Boutheina. 2022. "'A man is known by the company he keeps': A critical overview of semantic prosody." *Legon Journal of the Humanities* 33 (2): 59–79. <https://doi.org/10.4314/ljh.v33i2.3>.
- Channel, Joanna. 2000. "Corpus-based analysis of evaluative lexis." In *Evaluation in text: Authorial stance and the construction of ethos*, edited by Susan Hunston and Geoffrey Thompson, 38–55. Oxford University Press.
- Cheshire, Jenny. 2004. "Sex and gender in variationist research." In *The handbook of language variation and change*, edited by Jack K. Chambers, Peter Trudgill, and Natalie Schilling, 423–443. Blackwell.
- Coupland, Nikolas. 1997. "Language, ageing and ageism: A project for applied linguistics?" *International Journal of Applied Linguistics* 7 (1): 26–48. <https://doi.org/10.1111/j.1473-4192.1997.tb00102.x>.
- Coupland, Nikolas, Justine Coupland, and Howard Giles. 1991. *Language, society and the elderly. discourse, identity and aging*. Oxford, Cambridge: Blackwell.
- Eckert, Penelope. 1989. "The whole woman: Sex and gender differences in variation." *Language Variation and Change* 3 (1): 245–267.
- Escourrou, Émile, Sarah Laurent, Jacques Leroux, Stéphane Oustric, and Virginie Gardette. 2022. "The shift from old age to very old age: An analysis of the perception of aging among older people." *BMC Primary Care* 23 (3). <https://doi.org/10.1186/s12875-021-01616-4>.
- Firth, John Rupert. 1957. *Papers in linguistics 1934-1951*. London: Oxford University Press.
- Fulmer, Terry, David B. Reuben, John Auerbach, Donna Marie Fick, Colleen Galambos, and Kimberly S. Johnson. 2021. "Actualizing better health and health care for older adults." *Health Affairs* 40 (2): 219–225. <https://doi.org/10.1377/hlthaff.2020.01470>.

- Gerstenberg, Annette, and Camilla Lindholm. 2019. "Language and aging research: New insights and perspectives." *Linguistics Vanguard* 5 (2). <https://doi.org/10.1515/lingvan-2019-0025>.
- Hareven, Tamara. 1995. "Changing images of aging and the social construction of the life course." In *Images of aging: Cultural representations of later life*, edited by Mike Featherstone and Andrew Wernick, 119–134. Routledge.
- Hoey, Michael. 2005. *Lexical priming: A new theory of words and language*. London and New York: Routledge/AHRB.
- Hu, Hsueh-chao Marcella. 2015. "A semantic prosody analysis of three adjectives synonymous pairs in COCA." *Journal of Language and Linguistic Studies* 11 (2): 117–131.
- Hunston, Susan. 2007. "Semantic prosody revisited." *International Journal of Corpus Linguistics* 12 (2): 249–268.
- Jamet, Denis. 2018. "The neological functions of disease euphemisms in English and French: Verbal hygiene or speech pathology?" In *Lexis—Journal in English Lexicology 12: Lexical and semantic neology in English*, edited by Denis Jamet and Adeline Terry. <https://doi.org/10.4000/lexis.2397>.
- . 2022. "Les euphémismes de la maladie mentale : duplicités énonciatives ?" In *Quand le dire se dédouble : duplicité énonciative et stratégies discursives associées (domaine anglophone)*, edited by Blandine Pennec, 19–46. Toulouse: Presses Universitaires du Mirail.
- Jamet, Denis, and Christophe Coupé. 2023. "How is mental illness discursively constructed and conceptualized in a corpus of online blogs? A corpus-linguistics case-study of discourses on mental illness." *Anglophonia* 36. <http://journals.openedition.org/anglophonia/5605>.
- Jamet, Denis, and Bérengère Lafiandra. 2023. "'Crise migratoire' et discours politique d'extrême droite en France : 2015-2019. Analyse sémantique des discours de Marine Le Pen." *Mots. Les langages du politique* 131:119–144. <https://doi.org/10.4000/mots.30959>.
- Kilgarrieff, Adam, Vít Baisa, Jan Buta, Milo Jakubíek, Vojtěch Ková, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. "The Sketch Engine: Ten years on." *Lexicography* 1 (1): 7–36. <https://doi.org/10.1007/s40607-014-0009-9>.
- Louw, Bill. 1993. "Irony in the text or insincerity in the writer? The diagnosed potential of semantic prosodies." In *Text and technology: in honour of John Sinclair*, edited by Mona et al. Baker, 157–176. Benjamins.
- Mautner, Gerlinde. 2007. "Mining large corpora for social information: The case of elderly." *Language in Society* 36:51–72. <https://doi.org/10.1017/S0047404507070030>.
- Morley, John, and Alan Partington. 2009. "A few frequently asked questions about semantic—or evaluative—prosody." *International Journal of Corpus Linguistics* 14 (2): 139–158. <https://doi.org/10.1075/ijcl.14.2.01mor>.
- Nuessel, Frank H. 1982. "The language of ageism." *Gerontologist* 22:273–276.

- Nuessel, Frank H. 1984–1985. “Ageist language.” *Maledicta* 8:17–28.
- Partington, Alan. 2004. “‘Utterly content in each other’s company’: Semantic prosody and semantic preference.” *International Journal of Corpus Linguistics* 9:131–156. <https://doi.org/10.1075/ijcl.9.1.07par>.
- . 2015. “Evaluative prosody.” In *Corpus pragmatics: A handbook*, edited by Karen Aijmer and Rühlemann, Christoph, 279–303. Cambridge University Press.
- Rayson, Paul. 2009. *Wmatrix: A web-based corpus processing environment (Version 5)*. <http://ucrel.lancs.ac.uk/wmatrix/>.
- . 2023. *Free USAS English web tagger*. University of Lancaster. <http://ucrel-api.lancaster.ac.uk/usas/tagger.html>.
- Sinclair, John. 2004. *Trust the text: Language, corpus and discourse*. London and New York: Routledge.
- SketchEngine. 2022. *enTenTen: Corpus of the English web*. Brno: Lexical Computing. <https://www.sketchengine.eu/ententen-english-corpus/>.
- . 2017–2024. *SketchEngine*. Brno: Lexical Computing. <https://www.sketchengine.eu/>.
- Soto-Almela, Jorge, and Gema Alcaraz-Mármol. 2016. “The semantic prosody of the words inmigración and inmigrante in the Spanish written media: A corpus-based study of two national newspapers.” *Revista Signos: Estudios de Lingüística* 49 (91): 145–167.
- . 2017. “Comparing the portrayal of immigration in two Spanish newspapers: a semantic prosody analysis.” *Onomázein* 38:147–168. <https://doi.org/10.7764/onomazein.38.03>.
- Stewart, Dominic. 2010. *Semantic prosody: A critical evaluation*. London, England: Routledge.
- Stubbs, Michael. 2001. *Words and phrases*. Oxford: Blackwell.
- Tommaso, Laura. 2015. “The construction of age identity in an online discourse community: The case of *Boomer Women Speak*.” In *Language diversity: Identities, genres, discourses*, edited by G. Balirano and Nisco M. C., 163–175. Cambridge Scholars Publishing.
- Turner, Mark. 1998. *The literary mind*. Oxford University Press.
- Wangler, Julian, and Michael Jansky. 2023. “Media portrayal of old age and its effects on attitudes in older people: Findings from a series of studies.” *Humanities and Social Sciences Communications* 10:165. <https://doi.org/10.1057/s41599-023-01671-5>.
- Whitsitt, Sam. 2005. “A critique of the concept of semantic prosody.” *International Journal of Corpus Linguistics* 10 (3): 283–305. <https://doi.org/10.1075/ijcl.10.3.01whi>.

Xiao, Richard Z., and Anthony M. McEnery. 2006. "Collocation, semantic prosody and near synonymy: A cross-linguistic perspective." *Applied Linguistics* 27 (1): 103–129. <https://doi.org/10.1093/applin/ami045>.

Appendix

Presentation of enTenTen

The English Web Corpus (enTenTen) is an English corpus made up of texts collected from the Internet. The corpus belongs to the TenTen corpus family. Sketch Engine currently provides access to TenTen corpora in more than 40 languages. The corpora are built using technology specialized in collecting only linguistically valuable web content. For detailed information about TenTen corpora, see Common TenTen corpora attributes. The most recent version of the enTenTen corpus consists of 36 billion words. The texts were downloaded between 2019 and 2021. The sample texts of the biggest web domains which account for 40% of all corpus texts were checked semi-manually and content with poor quality text and spam was removed. (SketchEngine 2017–2024)

| | | | |
|--|--|---|--|
| GENERAL INFO Language: English CORPUS DESCRIPTION & BIBLIOGRAPHY TAGSET WORD SKETCH GRAMMAR TERM GRAMMAR | | COUNTS Tokens 43,125,207,462 Words 36,561,273,153 Sentences 2,008,143,278 Paragraphs 757,175,676 Documents 78,373,887 | LEXICON SIZES word? 96,177,617 tag 64 lemmas? 77,727,744 pos 10 lemma 72,590,429 lemmas_lc 67,276,299 lemma_lc 62,408,799 lc 84,864,356 |
|--|--|---|--|

Figure 5: General information on enTenTen 2020

Topic classification of the classified part of the English Web 2020 corpus

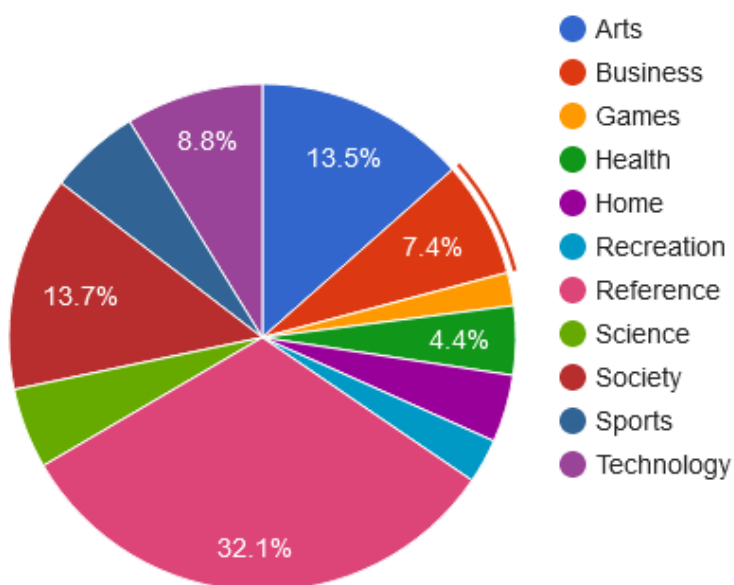


Figure 6: Topic classification in enTenTen 2020

| SUBCORPUS SIZES | | |
|------------------------|---------------|--------|
| Subcorpus | Tokens | % |
| Australian domain .au | 1,257,674,734 | 2.916 |
| Canadian domain .ca | 827,228,965 | 1.918 |
| EU domain .eu | 122,483 | 0.00 |
| English Wikipedia | 2,934,070,827 | 6.804 |
| Genre Blog | 278,159,566 | 0.645 |
| Genre Discussion | 312,926,515 | 0.726 |
| Genre Legal | 98,166,204 | 0.228 |
| Genre News | 1,560,229,567 | 3.618 |
| Irish domain .ie | 195,201,982 | 0.453 |
| New Zealand domain .nz | 288,926,496 | 0.67 |
| Topic Arts | 2,406,200,277 | 5.58 |
| Topic Business | 1,319,517,349 | 3.06 |
| Topic Games | 375,435,903 | 0.871 |
| Topic Health | 785,839,111 | 1.822 |
| Topic Home | 767,529,585 | 1.78 |
| Topic Recreation | 504,140,581 | 1.169 |
| Topic Reference | 5,718,808,057 | 13.261 |
| Topic Science | 915,297,931 | 2.122 |
| Topic Society | 2,447,004,798 | 5.674 |
| Topic Sports | 1,034,647,808 | 2.399 |
| Topic Technology | 1,566,713,474 | 3.633 |
| UK domain .uk | 3,420,336,927 | 7.931 |
| US domain .us | 382,681,188 | 0.887 |

Figure 7: Subcorpora in enTenTen 2020

CQL searches and normalized frequencies

Table 3: CQL searches and normalized frequencies for lemmatized and non-lemmatized tokens

| CQL searches for <i>non-lemmatized wordform</i> and LEMMA | Tokens | Normalized frequency (per million tokens) |
|---|---------|--|
| <i>old</i> MAN | 316 063 | 7.33 |
| <i>older people</i> | 134 325 | 3.11 |
| <i>old</i> WOMAN | 92 698 | 2.15 |
| <i>old woman</i> | 75 262 | 1.75 |
| <i>older</i> MAN | 64 754 | 1.5 |
| <i>older</i> WOMAN | 51 415 | 1.19 |
| <i>old people</i> | 41 532 | 0.96 |
| <i>elderly people</i> | 40 114 | 0.93 |
| <i>older man</i> | 37 486 | 0.87 |
| <i>old men</i> | 36 584 | 0.85 |
| <i>older women</i> | 27 351 | 0.63 |
| <i>elderly</i> WOMAN | 27 334 | 0.63 |
| <i>elderly</i> MAN | 24 453 | 0.57 |
| <i>older men</i> | 22 973 | 0.53 |
| <i>older woman</i> | 21 895 | 0.51 |
| <i>middle-aged</i> MAN | 16 036 | 0.37 |
| <i>old women</i> | 12 702 | 0.29 |
| <i>middle-aged</i> WOMAN | 10 746 | 0.25 |
| <i>middle-aged man</i> | 10 392 | 0.24 |
| <i>middle-aged woman</i> | 6 55 | 0.15 |
| <i>middle-aged men</i> | 5 405 | 0.13 |
| <i>mature</i> WOMAN | 4 754 | 0.11 |
| <i>middle-aged women</i> | 4 029 | 0.09 |
| <i>mature</i> MAN | 3 28 | 0.08 |
| <i>middle-aged people</i> | 2 043 | 0.05 |
| <i>mature people</i> | 1 706 | 0.04 |

Table 4: CQL searches and normalized frequencies for lemmatized tokens

| CQL search: LEMMA | Tokens | Normalized frequency (per million tokens) |
|---------------------------|---------|---|
| <i>old</i> MAN | 316 063 | 7.33 |
| <i>older people</i> | 134 325 | 3.11 |
| <i>old</i> WOMAN | 92 698 | 2.15 |
| <i>older</i> MAN | 64 754 | 1.5 |
| <i>older</i> WOMAN | 51 415 | 1.19 |
| <i>old people</i> | 41 532 | 0.96 |
| <i>elderly people</i> | 40 114 | 0.93 |
| <i>elderly</i> WOMAN | 27 334 | 0.63 |
| <i>elderly</i> MAN | 24 453 | 0.57 |
| <i>middle-aged</i> MAN | 16 036 | 0.37 |
| <i>middle-aged</i> WOMAN | 10 746 | 0.25 |
| <i>mature</i> WOMAN | 4 754 | 0.11 |
| <i>mature</i> MAN | 3 28 | 0.08 |
| <i>middle-aged people</i> | 2 043 | 0.05 |
| <i>mature people</i> | 1 706 | 0.04 |

Table 5: CQL searches and normalized frequencies for adjective + sequences to refer to AP (the tag **J.*** refers to adjectives)

| CQL search: tag , token, LEMMA | Tokens | Normalized frequency (per million tokens) |
|---------------------------------------|--------|---|
| J.* <i>old</i> MAN | 51 509 | 1.19 |
| J.* <i>old</i> WOMAN | 14 356 | 0.33 |
| J.* <i>older</i> MAN | 3 533 | 0.08 |
| J.* <i>older</i> WOMAN | 3 526 | 0.08 |
| J.* <i>middle-aged</i> MAN | 2 457 | 0.06 |
| J.* <i>elderly</i> MAN | 1 591 | 0.04 |
| J.* <i>elderly</i> WOMAN | 1 898 | 0.04 |
| J.* <i>middle-aged</i> WOMAN | 1 536 | 0.04 |
| J.* <i>mature</i> MAN | 254 | 0.01 |
| J.* <i>mature</i> WOMAN | 573 | 0.01 |

Semantic domains and adjectives used in attributive position

Table 6: Most frequent semantic domains of adjectives used in attributive position with (a) *old m?n* and (b) *old wom?n*

Labels of semantic domains: year_T1.3 dirty_O4.2- grumpy_E3- little_N3.2- wise_S1.2.6+ grand_O4.2+ poor_I1.1- good_A5.1+ crazy_B2-/X1 cranky_A6.2- bitter_X3.1 lonely_S5- frail_S1.2.5- strange_A6.2- creepy_E5- angry_E3- bearded_O4.2 fat_N3.2+ crotchety_Z99 nice_O4.2+ venerable_T3+ tired_B1 dear_E2+ feeble_S1.2.5- senile_T3++

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| E | Emotional actions, states and processes | 5 | 20 % |
| O | Substances, materials, objects and equipment | 5 | 20 % |
| S | Social actions, states and processes | 4 | 16 % |
| A | General and abstract terms | 3 | 12 % |
| T | Time | 3 | 12 % |
| N | Numbers and measurement | 2 | 8 % |
| B | The body and the individual | 1.5 | 6 % |
| I | Money and commerce in industry | 1 | 4 % |
| X | Psychological actions, states and processes | 0.5 | 2 % |
| | | 25 | 100 % |

(a)

Labels of semantic domains: year_T1.3 little_A13.7 poor_I1.1- wise_S1.2.6+ ugly_O4.2- lonely_S5- crazy_B2-/X1 bitter_X3.1 strange_A6.2- frail_S1.2.5- good_A5.1+ mysterious_A6.2- grumpy_E3- yr_Z99 silly_S1.2.6- eccentric_A6.2- fat_O1 bent_A2.1+ wicked_G2.2- rich_I1.1+ toothless_Z99 sweet_X3.1 sick_B2- mad_B2-/X1 wizened_Z99

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| A | General and abstract terms | 6 | 24 % |
| O | Substances, materials, objects and equipment | 4 | 16 % |
| S | Social actions, states and processes | 4 | 16 % |
| B | The body and the individual | 3 | 12 % |
| I | Money and commerce in industry | 2 | 8 % |
| T | Time | 2 | 8 % |
| E | Emotional actions, states and processes | 1 | 4 % |
| G | Govt. And the public domain | 1 | 4 % |
| N | Numbers and measurement | 1 | 4 % |
| X | Psychological actions, states and processes | 1 | 4 % |
| | | 25 | 100 % |

(b)

Table 7: Most frequent semantic domains of adjectives used in attributive position with (a) *older m?n* and (b) *older wom?n*

Labels of semantic domains: wealthy_I1.1+ healthy_B2+ rich_I1.1+ handsome_O4.2+ other_A6.1- attractive_O4.2+ powerful_S7.1+ white_O4.3 mysterious_A6.2- wise_S1.2.6+ lonely_S5- nice_O4.2+ haired_B1 sexy_S3.2 bearded_O4.2 distinguished_S7.1+ creepy_E5- predatory_L2 / F1 well_A5.1+ experienced_X2.2+ successful_X9.2+ tall_N3.7+ charismatic_S1.2 hot_O4.6+ charming_O4.2+ / A8[i1.2.1 looking_O4.2+ / A8[i1.2.2 married_S4 strange_A6.2-

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| S | Social actions, states and processes | 9 | 32.15 % |
| O | Substances, materials, objects and equipment | 7 | 25 % |
| A | General and abstract terms | 5 | 17.86 % |
| B | The body and the individual | 2 | 7.14 % |
| I | Money and commerce in industry | 2 | 7.14 % |
| X | Psychological actions, states and processes | 2 | 7.14 % |
| N | Numbers and measurement | 1 | 3.57 % |
| | | 25 | 100 % |

(a)

Labels of semantic domains: other_A6.1- attractive_O4.2+ beautiful_O4.2+ wealthy_I1.1+ healthy_B2+ sexy_S3.2 single_N5— rich_I1.1+ hot_O4.6+ frail_S1.2.5- mysterious_A6.2- nice_O4.2+ sweet_X3.1 lonely_S5- wise_S1.2.6+ lovely_O4.2+ well_A5.1+ glamorous_O4.2+ elegant_O4.2+ experienced_X2.2+ sophisticated_O4.2+ eccentric_A6.2- haired_B1 powerful_S7.1+ seductive_S3.2

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| O | Substances, materials, objects and equipment | 8 | 32 % |
| S | Social actions, states and processes | 8 | 32 % |
| A | General and abstract terms | 4 | 16 % |
| B | The body and the individual | 2 | 8 % |
| I | Money and commerce in industry | 2 | 8 % |
| X | Psychological actions, states and processes | 1 | 4 % |
| | | 25 | 100 % |

(b)

Table 8: Most frequent semantic domains of adjectives used in attributive position with (a) *elderly m?n* and (b) *elderly wom?n*

Labels of semantic domains: healthy_B2+ frail_S1.2.5- / A8[i1.2.1 looking_S1.2.5- / A8[i1.2.2 other_A6.1- missing_A3- unidentified_X2.2- wealthy_I1.1+ old_T3+ small_N3.2- disabled_B2- nice_O4.2+ same_A6.1+++ vulnerable_S1.2.5- little_N3.2- poor_I1.1- single_N5— local_M7 sick_B2- stout_O4.2 bearded_O4.2 distinguished_S7.1+ haired_B1 well_W3 / M4

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| A | General and abstract terms | 5 | 21.73 % |
| B | The body and the individual | 4 | 17.39 % |
| S | Social actions, states and processes | 4 | 17.39 % |
| O | Substances, materials, objects and equipment | 3 | 13.04 % |
| I | Money and commerce in industry | 2 | 8.70 % |
| N | Numbers and measurement | 2 | 8.70 % |
| M | Movement, location, travel and transport | 1 | 4.35 % |
| T | Time | 1 | 4.35 % |
| X | Psychological actions, states and processes | 1 | 4.35 % |
| | | 23 | 100 % |

(a)

Labels of semantic domains: frail_S1.2.5- other_A6.1- poor_I1.1- single_N5— sweet_X3.1 healthy_B2+ local_M7 small_N3.2- vulnerable_S1.2.5- wealthy_I1.1+ missing_A3- nice_O4.2+ old_T3+ / A8[i1.2.1 looking_T3+ / A8[i1.2.2 little_A13.7 haired_B1 same_A6.1+++ confused_X2.5- eccentric_A6.2- ill_B2- sick_B2- unidentified_X2.2- disabled_B2- year_T1.3

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| A | General and abstract terms | 6 | 25 % |
| B | The body and the individual | 5 | 20.84 % |
| S | Social actions, states and processes | 3 | 12.50 % |
| I | Money and commerce in industry | 2 | 8.33 % |
| O | Substances, materials, objects and equipment | 2 | 8.33 % |
| T | Time | 2 | 8.33 % |
| X | Psychological actions, states and processes | 2 | 8.33 % |
| M | Movement, location, travel and transport | 1 | 4.17 % |
| N | Numbers and measurement | 1 | 4.17 % |
| | | 25 | 100 % |

(b)

Table 9: Most frequent semantic domains of adjectives used in attributive position with (a) *middle-aged m?n* and (b) *middle-aged wom?n*

Labels of semantic domains: healthy_B2+ white_O4.3 looking_B1 / A8[i1.2.1 balding_B1 / A8[i1.2.2 overweight_N3.5 / N5.2+ ordinary_A6.2+ portly_O4.2 late_T4- well_A5.1+ angry_E3- fat_N3.2+ hand-some_O4.2+ lonely_S5- other_A6.1- dressed_B5 bearded_O4.2 obese_N3.4+ tall_N3.7+ average_A6.2+ wealthy_I1.1+ married_S4 stout_O4.2

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| A | General and abstract terms | 5 | 22.72 % |
| O | Substances, materials, objects and equipment | 5 | 22.72 % |
| N | Numbers and measurement | 4 | 18.18 % |
| B | The body and the individual | 3 | 13.64 % |
| S | Social actions, states and processes | 2 | 9.09 % |
| E | Emotional actions, states and processes | 1 | 4.55 % |
| I | Money and commerce in industry | 1 | 4.55 % |
| T | Time | 1 | 4.55 % |
| | | 22 | 100 % |

(a)

Labels of semantic domains: attractive_O4.2+ / A8[i1.2.1 looking_O4.2+ / A8[i1.2.2 white_O4.3 healthy_B2+ lonely_S5- other_A6.1- plump_O4.4 obese_N3.4+ overweight_N3.5 / N5.2+ stout_O4.2 well_A5.1+ beautiful_O4.2+ respectable_G2.2+ late_T4- pleasant_O4.2+ single_N5— ordinary_A6.2+ plain_O4.1 wealthy_I1.1+ dressed_B5 handsome_O4.2+

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| O | Substances, materials, objects and equipment | 8 | 38.11 % |
| A | General and abstract terms | 4 | 19.05 % |
| B | The body and the individual | 2 | 9.52 % |
| N | Numbers and measurement | 2 | 9.52 % |
| S | Social actions, states and processes | 2 | 9.52 % |
| G | Govt. And the public domain | 1 | 4.76 % |
| I | Money and commerce in industry | 1 | 4.76 % |
| T | Time | 1 | 4.76 % |
| | | 21 | 100 % |

(b)

Table 10: Most frequent semantic domains of adjectives used in attributive position with (a) *mature m?n* and (b) *mature wom?n*

Labels of semantic domains: older_T3++ real_A3+ serious_A11.1+ loving_E2+ other_A6.1- strong_S1.2.5+ young_T3- active_X5.2+ hairy_O4.1 / O4.5 healthy_B2+ intelligent_X9.1+ old_T3+ qualified_A1.2+ / I3.2+ single_N5— small_N3.2- affectively_Z99 bisexual_S3.2 / B1 caring_E2+ confident_E6+ decent_A5.1+ elegant_O4.2+ gay_S3.2 / B1 horny_Z99 hung_M2 little_A13.7 / E2+[i1.2.1 minded_A13.7 / E2+[i1.2.2 sexy_S3.2 straight_O4.4

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| S | Social actions, states and processes | 7 | 25 % |
| A | General and abstract terms | 5.5 | 19.65 % |
| E | Emotional actions, states and processes | 5 | 17.86 % |
| T | Time | 3 | 10.71 % |
| O | Substances, materials, objects and equipment | 2 | 7.14 % |
| X | Psychological actions, states and processes | 2 | 7.14 % |
| B | The body and the individual | 1 | 3.57 % |
| M | Movement, location, travel and transport | 1 | 3.57 % |
| N | Numbers and measurement | 1 | 3.57 % |
| I | Money and commerce in industry | 0.5 | 1.79 % |
| | | 28 | 100 % |

(a)

Labels of semantic domains: beautiful_O4.2+ sexy_S3.2 hot_O4.6+ older_T3++ attractive_O4.2+ old_T3+ horny_Z99 other_A6.1- single_N5— experienced_X2.2+ nude_B5- busty_Z99 elegant_O4.2+ gorgeous_O4.2+ lovely_O4.2+ chubby_O4.2 normal_A6.2+ sophisticated_O4.2+ strong_S1.2.5+ wild_L1

| Code | Semantic domain | Tokens | Percent |
|------|--|--------|---------|
| O | Substances, materials, objects and equipment | 8 | 40 % |
| S | Social actions, states and processes | 5 | 25 % |
| A | General and abstract terms | 2 | 10 % |
| T | Time | 2 | 10 % |
| B | The body and the individual | 1 | 5 % |
| L | Life and living things | 1 | 5 % |
| X | Psychological actions, states and processes | 1 | 5 % |
| | | 20 | 100 % |

(b)

Most frequent adjectives used in attributive position

Table 11: Most frequent adjectives used in attributive position with (a) *old m?n* and (b) *old wom?n* (values in *italics* are manually corrected)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|------------------|-----------|------|------|--------------|--|
| year | 67162.737 | 1 | 695 | T1.3 | Time: Period |
| dirty | 47158.871 | 2 | 488 | O4.2- | Judgement of appearance (pretty etc.) |
| grumpy | 38364.901 | 3 | 397 | E3- | Calm / Violent / Angry |
| little | 38074.990 | 4 | 394 | N3.2- | Measurement: Size |
| wise | 31793.583 | 5 | 329 | S1.2.6+ | Sensible |
| grand | 31696.946 | 6 | 328 | O4.2+ | Judgement of appearance (pretty etc.) |
| poor | 25705.450 | 7 | 266 | I1.1- | Money: Affluence |
| good | 18361.036 | 8 | 190 | A5.1+ | Evaluation: Good / bad |
| crazy | 13239.273 | 9 | 137 | B2- / X1 | Health and disease / General |
| cranky | 11886.355 | 10 | 123 | A6.2- | Comparing: Usual / unusual |
| bitter | 11306.533 | 11 | 117 | E2- | Liking |
| lonely | 10630.073 | 12 | 110 | S5- | Groups and affiliation |
| frail | 10146.888 | 13 | 105 | S1.2.5- | Toughness; strong / weak |
| strange | 8697.333 | 14 | 90 | A6.2- | Comparing: Usual / unusual |
| creepy | 8504.059 | 15 | 88 | O4.2- | Judgement of appearance (pretty etc.) |
| angry | 8214.148 | 16 | 85 | E3- | Calm / Violent / Angry |
| bearded | 8117.511 | 17 | 84 | O4.2 | Judgement of appearance (pretty etc.) |
| fat | 7924.237 | 18 | 82 | N3.5 / N5.2+ | Measurement: Weight / Exceeding; waste |
| crotchety | 7730.963 | 19 | 80 | E3- | Calm / Violent / Angry |
| nice | 7730.963 | 19 | 80 | O4.2+ | Judgement of appearance (pretty etc.) |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|-------------------|------------|------|------|--------------|--|
| year | 194935.730 | 1 | 2017 | T1.3 | Time: Old, new and young; age |
| little | 59824.104 | 2 | 619 | A13.7 | Degree: Minimizers |
| poor | 50739.345 | 3 | 525 | I1.1- | Money: Affluence |
| wise | 19715.860 | 4 | 204 | S1.2.6+ | Sensible |
| ugly | 14303.663 | 5 | 148 | O4.2- | Judgement of appearance (pretty etc.) |
| lonely | 14207.017 | 6 | 147 | S5- | Groups and affiliation |
| crazy | 10727.747 | 7 | 111 | B2- / X1 | Health and disease / General |
| bitter | 9567.991 | 8 | 99 | O4.2 | Judgement of appearance (pretty etc.) |
| strange | 9374.698 | 9 | 97 | A6.2- | Comparing: Usual / unusual |
| frail | 9181.405 | 10 | 95 | S1.2.5- | Toughness; strong / weak |
| good | 9181.405 | 10 | 95 | A5.1+ | Evaluation: Good / bad |
| mysterious | 8311.588 | 12 | 86 | A6.2- | Comparing: Usual / unusual |
| grumpy | 7345.124 | 13 | 76 | E3- | Calm / Violent / Angry |
| yr | 6958.539 | 14 | 72 | T1.3 | Time: Old, new and young; age |
| silly | 6861.892 | 15 | 71 | S1.2.6- | Sensible |
| eccentric | 6571.953 | 16 | 68 | A6.2- | Comparing: Usual / unusual |
| fat | 6571.953 | 16 | 68 | N3.5 / N5.2+ | Measurement: Weight / Exceeding; waste |
| bent | 6378.660 | 18 | 66 | A2.1+ | Affect: Modify, change |
| wicked | 6185.368 | 19 | 64 | G2.2- | General ethics |
| rich | 5605.490 | 20 | 58 | I1.1+ | Money: Affluence |
| toothless | 5605.490 | 20 | 58 | B1 | Anatomy and physiology |

(b)

Table 12: Most frequent adjectives used in attributive position with (a) *older m?n* and (b) *older wom?n* (values in *italics* are manually corrected)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|----------------------|-----------|------|------|---------|---------------------------------------|
| wealthy | 49939.099 | 1 | 164 | I1.1+ | Money: Affluence |
| healthy | 34409.257 | 2 | 113 | B2+ | Health and disease |
| rich | 29537.150 | 3 | 97 | I1.1+ | Money: Affluence |
| handsome | 28623.630 | 4 | 94 | O4.2+ | Judgement of appearance (pretty etc.) |
| other | 26796.590 | 5 | 88 | A6.1- | Comparing: Similar / different |
| attractive | 14311.815 | 6 | 47 | O4.2+ | Judgement of appearance (pretty etc.) |
| powerful | 14311.815 | 6 | 47 | S7.1+ | Power, organizing |
| white | 13398.295 | 8 | 44 | O4.3 | Color and color patterns |
| mysterious | 11266.748 | 9 | 37 | A6.2- | Comparing: Usual / unusual |
| wise | 10962.241 | 10 | 36 | S1.2.6+ | Sensible |
| lonely | 10353.228 | 11 | 34 | S5- | Groups and affiliation |
| nice | 10353.228 | 11 | 34 | O4.2+ | Judgement of appearance (pretty etc.) |
| haired | 9439.708 | 13 | 31 | B1 | Anatomy and physiology |
| sexy | 9135.201 | 14 | 30 | S3.2 | Relationship: Intimate / sexual |
| bearded | 8830.694 | 15 | 29 | O4.2 | Judgement of appearance (pretty etc.) |
| distinguished | 8526.188 | 16 | 28 | S7.1+ | Power, organizing |
| creepy | 8221.681 | 17 | 27 | O4.2- | Judgement of appearance (pretty etc.) |
| predatory | 8221.681 | 17 | 27 | S1.2- | Personality traits |
| well | 8221.681 | 17 | 27 | A5.1+ | Evaluation: Good / bad |
| experienced | 7917.174 | 20 | 26 | X2.2+ | Knowledge |
| successful | 7917.174 | 20 | 26 | X9.2+ | Ability: Success and failure |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|----------------------|-----------|------|------|---------|---------------------------------|
| other | 41051.956 | 1 | 128 | A6.1- | Comparing: Similar / different |
| attractive | 36882.617 | 2 | 115 | O4.2+ | Judgement of appearance |
| beautiful | 34637.588 | 3 | 108 | O4.2+ | Judgement of appearance |
| wealthy | 27261.065 | 4 | 85 | I1.1+ | Money: Affluence |
| healthy | 24374.599 | 5 | 76 | B2+ | Health and disease |
| sexy | 23091.725 | 6 | 72 | S3.2 | Relationship: Intimate / sexual |
| single | 20846.697 | 7 | 65 | S4 | Kin |
| rich | 16356.639 | 8 | 51 | I1.1+ | Money: Affluence |
| hot | 14111.610 | 9 | 44 | S3.2 | Relationship: Intimate / sexual |
| frail | 11225.144 | 10 | 35 | S1.2.5- | Toughness; strong / weak |
| mysterious | 11225.144 | 10 | 35 | A6.2- | Comparing: Usual / unusual |
| nice | 9942.271 | 12 | 31 | O4.2+ | Judgement of appearance |
| sweet | 9942.271 | 12 | 31 | O4.2+ | Judgement of appearance |
| lonely | 9300.834 | 14 | 29 | S5- | Groups and affiliation |
| wise | 8980.115 | 15 | 28 | S1.2.6+ | Sensible |
| lovely | 8338.679 | 16 | 26 | O4.2+ | Physical attributes |
| well | 8338.679 | 16 | 26 | A5.1+ | Evaluation: Good / bad |
| glamorous | 8017.960 | 18 | 25 | O4.2+ | Judgement of appearance |
| elegant | 7697.242 | 19 | 24 | O4.2+ | Judgement of appearance |
| experienced | 7697.242 | 19 | 24 | X2.2+ | Knowledge |
| sophisticated | 7697.242 | 19 | 24 | O4.2+ | Judgement of appearance |

(b)

Table 13: Most frequent adjectives used in attributive position with (a) *elderly m?n* and (b) *elderly wom?n* (values in *italics* are manually corrected)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|----------------------|-----------|------|------|----------------|---------------------------------------|
| healthy | 34629.405 | 1 | 57 | B2+ | Health and disease |
| frail | 21871.203 | 2 | 36 | <i>S1.2.5-</i> | Toughness; strong / weak |
| looking | 21871.203 | 2 | 36 | A8 | Seem / Appear |
| other | 20048.603 | 4 | 33 | A6.1- | Comparing: Similar / different |
| missing | 18833.536 | 5 | 31 | A3- | Being |
| unidentified | 14580.802 | 6 | 24 | X2.2- | Knowledge |
| wealthy | 14580.802 | 6 | 24 | I1.1+ | Money: Affluence |
| old | 13973.269 | 8 | 23 | T3+ | Time: Old, new and young; age |
| small | 13365.735 | 9 | 22 | N3.2- | Measurement: Size |
| disabled | 10935.601 | 10 | 18 | B2- | Health and disease |
| nice | 10935.601 | 10 | 18 | O4.2+ | Judgement of appearance (pretty etc.) |
| same | 10935.601 | 10 | 18 | A6.1+++ | Comparing: Similar / different |
| vulnerable | 10935.601 | 10 | 18 | <i>S1.2.5-</i> | Toughness; strong / weak |
| little | 10328.068 | 14 | 17 | N3.2- | Measurement: Size |
| poor | 10328.068 | 14 | 17 | I1.1- | Money: Affluence |
| single | 10328.068 | 14 | 17 | S4 | Kin |
| local | 9720.535 | 17 | 16 | M7 | Places |
| sick | 9720.535 | 17 | 16 | B2- | Health and disease |
| stout | 9720.535 | 17 | 16 | O4.2 | Judgement of appearance (pretty etc.) |
| bearded | 9113.001 | 20 | 15 | O4.2 | Judgement of appearance (pretty etc.) |
| distinguished | 9113.001 | 20 | 15 | S7.1+ | Power, organizing |
| haired | 9113.001 | 20 | 15 | B1 | Anatomy and physiology |
| well | 9113.001 | 20 | 15 | <i>A5.1+</i> | Evaluation: Good / bad |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|---------------------|-----------|------|------|----------------|--------------------------------|
| frail | 35002.693 | 1 | 65 | <i>S1.2.5-</i> | Toughness; strong / weak |
| other | 33925.687 | 2 | 63 | A6.1- | Comparing: Similar / different |
| poor | 20463.113 | 3 | 38 | I1.1- | Money: Affluence |
| single | 20463.113 | 3 | 38 | S4 | Kin |
| sweet | 16693.592 | 5 | 31 | O4.2+ | Physical attributes |
| healthy | 16155.089 | 6 | 30 | B2+ | Health and disease |
| local | 15616.586 | 7 | 29 | M7 | Places |
| small | 15078.083 | 8 | 28 | N3.2- | Measurement: Size |
| vulnerable | 14539.580 | 9 | 27 | <i>S1.2.5-</i> | Toughness; strong / weak |
| wealthy | 14001.077 | 10 | 26 | I1.1+ | Money: Affluence |
| missing | 13462.574 | 11 | 25 | A3- | Being |
| nice | 12924.071 | 12 | 24 | O4.2+ | Physical attributes |
| old | 11308.562 | 13 | 21 | T3+ | Time: Old, new and young; age |
| looking | 10770.059 | 14 | 20 | A8 | Seem / Appear |
| little | 10231.556 | 15 | 19 | A13.7 | Degree: Minimizers |
| haired | 9154.550 | 16 | 17 | B1 | Anatomy and physiology |
| same | 9154.550 | 16 | 17 | A6.1+++ | Comparing: Similar / different |
| confused | 8616.047 | 18 | 16 | X2.5- | Understand |
| eccentric | 8616.047 | 18 | 16 | A6.2- | Comparing: Usual / unusua |
| ill | 8616.047 | 18 | 16 | B2- | Health and disease |
| sick | 8616.047 | 18 | 16 | B2- | Health and disease |
| unidentified | 8616.047 | 18 | 16 | X2.2- | Knowledge |

(b)

Table 14: Most frequent adjectives used in attributive position with (a) *middle-aged m?n* and (b) *middle-aged wom?n* (values in *italics* are manually corrected)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|-------------------|-----------|------|------|--------------|--|
| healthy | 36287.242 | 1 | 95 | B2+ | Health and disease |
| white | 31321.620 | 2 | 82 | O4.3 | Color and color patterns |
| looking | 26737.968 | 3 | 70 | A8 | Seem / Appear |
| balding | 23682.200 | 4 | 62 | B1 | Anatomy and physiology |
| overweight | 23682.200 | 4 | 62 | N3.5 / N5.2+ | Measurement: Weight / Exceeding; waste |
| ordinary | 12605.042 | 6 | 33 | A6.2+ | Comparing: Usual / unusual |
| portly | 12605.042 | 6 | 33 | O4.2+ | Judgement of appearance (pretty etc.) |
| late | 12223.071 | 8 | 32 | T4- | Time: Early / late |
| well | 11841.100 | 9 | 31 | A5.1+ | Evaluation: Good / bad |
| angry | 10695.187 | 10 | 28 | E3- | Calm / Violent / Angry |
| fat | 10695.187 | 10 | 28 | N3.5 / N5.2+ | Measurement: Weight / Exceeding; waste |
| handsome | 10313.216 | 12 | 27 | O4.2+ | Judgement of appearance (pretty etc.) |
| lonely | 9549.274 | 13 | 25 | S5- | Groups and affiliation |
| other | 9167.303 | 14 | 24 | A6.1- | Comparing: Similar / different |
| dressed | 8403.361 | 15 | 22 | B5 | Clothes and personal belongings |
| bearded | 7639.419 | 16 | 20 | O4.2+ | Judgement of appearance (pretty etc.) |
| obese | 7639.419 | 16 | 20 | N3.5 / N5.2+ | Measurement: Weight / Exceeding; waste |
| tall | 7639.419 | 16 | 20 | N3.7+ | Measurement: Length & height |
| average | 7257.448 | 19 | 19 | A6.2+ | Comparing: Usual / unusual |
| wealthy | 7257.448 | 19 | 19 | I1.1+ | Money: Affluence |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|--------------------|-----------|------|------|-------------|--|
| attractive | 30978.934 | 1 | 50 | O4.2+ | Judgement of appearance (pretty etc.) |
| looking | 28500.620 | 2 | 46 | A8 | Seem / Appear |
| white | 24163.569 | 3 | 39 | O4.3 | Color and color patterns |
| healthy | 23543.990 | 4 | 38 | B2+ | Health and disease |
| lonely | 19206.939 | 5 | 31 | S5- | Groups and affiliation |
| other | 14869.888 | 6 | 24 | A6.1- | Comparing: Similar / different |
| plump | 14869.888 | 6 | 24 | O4.4 | Shape |
| obese | 12391.574 | 8 | 20 | N3.5 / N5.2 | Measurement: Weight / Exceeding; waste |
| overweight | 12391.574 | 8 | 20 | N3.5 / N5.2 | Measurement: Weight / Exceeding; waste |
| stout | 12391.574 | 8 | 20 | O4.2 | Judgement of appearance (pretty etc.) |
| well | 12391.574 | 8 | 20 | A5.1+ | Evaluation: Good / bad |
| beautiful | 11152.416 | 12 | 18 | O4.2+ | Judgement of appearance (pretty etc.) |
| respectable | 10532.838 | 13 | 17 | G2.2+ | General ethics |
| late | 9293.680 | 14 | 15 | T4- | Time: Early / late |
| pleasant | 8674.102 | 15 | 14 | O4.2+ | Judgement of appearance (pretty etc.) |
| single | 8674.102 | 15 | 14 | S4 | Kin |
| ordinary | 8054.523 | 17 | 13 | A6.2+ | Comparing: Usual / unusual |
| plain | 8054.523 | 17 | 13 | O4.1 | General appearance and physical properties |
| wealthy | 7434.944 | 19 | 12 | I1.1+ | Money: Affluence |
| dressed | 6815.366 | 20 | 11 | B5 | Clothes and personal belongings |
| handsome | 6815.366 | 20 | 11 | O4.2+ | Judgement of appearance (pretty etc.) |

(b)

Table 15: Most frequent adjectives used in attributive position with (a) *mature m?n* and (b) *mature wom?n* (values in *italics* are manually corrected)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|--------------------|-----------|------|------|---------------|--|
| older | 79096.045 | 1 | 14 | T3++ | Time: Old, new and young; age |
| real | 39548.023 | 2 | 7 | A3+ | Being |
| serious | 28248.588 | 3 | 5 | A11.1+ | Importance: Important |
| loving | 22598.870 | 4 | 4 | E2+ | Liking |
| other | 22598.870 | 4 | 4 | A6.1- | Comparing: Similar / different |
| strong | 22598.870 | 4 | 4 | S1.2.5+ | Toughness; strong / weak |
| young | 22598.870 | 4 | 4 | T3- | Time: Old, new and young; age |
| active | 16949.153 | 8 | 3 | X5.2+ | Interest / boredom / excited / energetic |
| hairy | 16949.153 | 8 | 3 | O4.1 | General appearance and physical properties |
| healthy | 16949.153 | 8 | 3 | B2+ | Health and disease |
| intelligent | 16949.153 | 8 | 3 | X9.1+ | Ability: Ability, intelligence |
| old | 16949.153 | 8 | 3 | T3+ | Time: Old, new and young; age |
| qualified | 16949.153 | 8 | 3 | A1.2+ / I3.2+ | Suitability / Work and employment: Professionalism |
| single | 16949.153 | 8 | 3 | S4 | Kin |
| small | 16949.153 | 8 | 3 | N3.2- | Measurement: Size |
| affectively | 11299.435 | 16 | 2 | E2+ | Liking |
| bisexual | 11299.435 | 16 | 2 | S3.2 | Relationship: Intimate / sexual |
| caring | 11299.435 | 16 | 2 | E2+ | Liking |
| confident | 11299.435 | 16 | 2 | E6+ | Worry, concern, confident |
| decent | 11299.435 | 16 | 2 | A5.1+ | Evaluation: Good / bad |
| elegant | 11299.435 | 16 | 2 | O4.2+ | Judgement of appearance (pretty etc.) |
| gay | 11299.435 | 16 | 2 | S3.2 | Relationship: Intimate / sexual |
| horny | 11299.435 | 16 | 2 | S3.2 | Relationship: Intimate / sexual |
| hung | 11299.435 | 16 | 2 | M2 | Putting, taking, pulling, pushing |
| little | 11299.435 | 16 | 2 | A13.7 | Degree: Minimizers |
| minded | 11299.435 | 16 | 2 | E6+ | Worry, concern, confident |
| sexy | 11299.435 | 16 | 2 | S3.2 | Relationship: Intimate / sexual |
| straight | 11299.435 | 16 | 2 | S3.2 | Relationship: Intimate / sexual |

(a)

| Type | NormFreq | Rank | Freq | Semtag | Semantic domain |
|----------------------|-----------|------|------|---------|---------------------------------------|
| beautiful | 92857.143 | 1 | 39 | O4.2+ | Judgement of appearance (pretty etc.) |
| sexy | 66666.667 | 2 | 28 | S3.2 | Relationship: Intimate / sexual |
| hot | 47619.048 | 3 | 20 | S3.2 | Relationship: Intimate / sexual |
| older | 33333.333 | 4 | 14 | T3++ | Time: Old, new and young; age |
| attractive | 30952.381 | 5 | 13 | O4.2+ | Judgement of appearance (pretty etc.) |
| old | 21428.571 | 6 | 9 | T3+ | Time: Old, new and young; age |
| horny | 19047.619 | 7 | 8 | S3.2 | Relationship: Intimate / sexual |
| other | 19047.619 | 7 | 8 | A6.1- | Comparing: Similar / different |
| single | 19047.619 | 7 | 8 | S4 | Kin |
| experienced | 14285.714 | 10 | 6 | X2.2+ | Knowledge |
| nude | 14285.714 | 10 | 6 | B5- | Clothes and personal belongings |
| busty | 11904.762 | 12 | 5 | O4.2+ | Judgement of appearance (pretty etc.) |
| elegant | 11904.762 | 12 | 5 | O4.2+ | Judgement of appearance (pretty etc.) |
| gorgeous | 11904.762 | 12 | 5 | O4.2+ | Judgement of appearance (pretty etc.) |
| lovely | 11904.762 | 12 | 5 | O4.2+ | Judgement of appearance (pretty etc.) |
| chubby | 9523.810 | 16 | 4 | O4.2 | Judgement of appearance (pretty etc.) |
| normal | 9523.810 | 16 | 4 | A6.2+ | Comparing: Usual / unusual |
| sophisticated | 9523.810 | 16 | 4 | O4.2+ | Judgement of appearance (pretty etc.) |
| strong | 9523.810 | 16 | 4 | S1.2.5+ | Toughness; strong / weak |
| wild | 9523.810 | 16 | 4 | L1 | Life and living things |

(b)

Word cloud visualization



Figure 8: Most frequent adjectives used in attributive position with (a) *old m?n* and (b) *old wom?n*



Figure 9: Most frequent adjectives used in attributive position with (a) *older m?n* and (b) *older wom?n*



Figure 10: Most frequent adjectives used in attributive position with (a) *elderly m?n* and (b) *elderly wom?n*



Figure 11: Most frequent adjectives used in attributive position with (a) *middle-aged m?n* and (b) *middle-aged wom?n*



Figure 12: Most frequent semantic domains of adjectives used in attributive position with (a) *mature m?n* and (b) *mature wom?n*