

Leeds Studies in English

New Series XLVII

2016

Edited by

Alaric Hall



Reviews editor
N. Kivilcim Yavuz

Leeds Studies in English

<www.leeds.ac.uk/lse>

School of English
University of Leeds

2016

Diachronic Development of the Order of Prenominal Adjectives in English: The Case of AGE and SHAPE Semantic Categories

Lukasz Stolarski

Introduction

In English there may be several adjectives in the attributive position and their order seems to follow particular patterns. For instance, the phrase ‘a funny red hat’ sounds more natural than ‘a red funny hat’ and the expression ‘an intriguing, small, round, yellow toy’ is more appropriate than ‘a yellow, round, small, intriguing toy’. Because of such tendencies to favour one pattern over others, numerous grammar books for foreign learners of English propose ready solutions on the order of adjectives in the attributive position.¹ In many such publications it is suggested that the patterns are only strong tendencies and alternative orders may also be found in English. Moreover, many details of particular solutions differ from each other and, ultimately, it may be difficult for a foreigner to apply the patterns in practice. Suggestions on particular orders are also presented in the scholarly linguistic literature. A summary of the adjective patterns proposed in a selection of publications is provided in Table 1.

The phenomenon under discussion has been explained by referring to distinctions along ‘general – specific’, ‘extrinsic – intrinsic’ or ‘subjective – objective’ continua.² To give a few

¹ For example, Louis Alexander, *Longman English Grammar Practice for Intermediate Students* (Harlow: Longman, 1996); Virginia Evans, *CPE Use of English* (Newbury: Express Publishing, 2008) and *FCE Use of English* (Newbury: Express Publishing, 2010); Martin Hewings, *Advanced Grammar in Use* (Cambridge: Cambridge University Press, 2002); J. Hill, R. Hurst, M. Lewis, C. Blissett and C. Hallgarten, *Grammar and Practice* (Hove: Language Teaching Publications, 1995); N. Hopkins and D. Hopkins, *Developing Grammar in Context* (Cambridge: Cambridge University Press, 2009); E. Mańczak-Wohlfeld, A. Niżegorodcew, and E. Willim, *A Practical Grammar of English* (Warsaw: Wydawnictwo Naukowe PWN, 1996); Martin Parrott, *Grammar for English Language Teachers* (Cambridge: Cambridge University Press, 2011); Fiona Scott-Barrett, *New Proficiency Use of English* (Harlow: Longman, 2002); Michael Swan, *Practical English Usage* (Oxford: Oxford University Press, 1995); George Yule, *Oxford Practice Grammar* (Oxford: Oxford University Press, 2011).

² Cf. Brian Byrne, ‘Rules of Prenominal Adjective Order and the Interpretation of “Incompatible” Adjective Pairs’, *Journal of Verbal Learning and Verbal Behavior*, 18 (1979), 73–78; J. H. Danks and S. Glucksberg, ‘Psychological Scaling of Adjective Orders’, *Journal of Verbal Learning & Verbal Behavior*, 1.10 (1971), 63–67; William Frawley, *Linguistic Semantics* (Hillsdale: Lawrence Erlbaum, 1992); Robert Hetzron, ‘On the Relative Order of Adjectives’, in *Language Universals*, ed. by Hans Sauer (Tubingen: Narr, 1978), pp. 165–84; Hill and others, *Grammar and Practice*; James Martin, ‘Semantic Determinants of Preferred Adjective Order’, *Journal of Verbal Learning and Verbal Behavior*, 8 (1969), 697–704 and ‘Some Competence-Process Relationships in

Diachronic Development of the Order of Prenominal Adjectives in English

examples, Hill and others claim that ‘usually the more specific the adjective is, the closer it comes to the noun. In other words, the order is often: general adjective – specific adjective – noun’. However, Teschner and Evans suggest that ‘the more intrinsic the adjective is to the nature of the noun, the closer it will be to the noun’.³ Such proposals are discussed by Danks and Glucksberg, who suggest an additional interpretation. They claim that adjectives which are more intrinsic to a modified noun are less likely to discriminate the noun referent from other potential referents. As a consequence, such adjectives tend to appear close to the noun. Conversely, adjectives which are less intrinsic to a modified noun are more suitable for discriminating the noun referent from other referents and this is why they tend to appear further away from the noun. For instance, in the phrase ‘a large red car’ the word ‘red’ is closer to the head than ‘large’ because it may easily be understood without reference to other objects. The notion of COLOUR tends to be more inherent to the noun ‘car’ than the notion of SIZE because the latter is more relative and a comparison to other cars is necessary for correct interpretation. Therefore, adjectives denoting SIZE are frequently more appropriate for discriminating objects and are placed further away from nouns than adjectives referring to COLOUR. Obviously, this may change in a situation in which one is referring to several large vehicles and wants to discriminate one of them by their colour. In such a case the more appropriate order would be ‘a red large car’.⁴

Other explanations involve various systems based on ‘zones’.⁵ In general, it is claimed that the position of a particular adjective within a noun phrase depends on which ‘modification zone’ it is placed in. Three of these are usually distinguished. The first one involves specifying adjectives which ‘help single out or quantify the referent of the construction in relation to some context’.⁶ They tend to have determiner-like properties. An example of a specifying adjective is ‘main’ in ‘his main reason’ and ‘former’ in ‘my former colleague’. Such items are placed furthest from the noun, although, obviously, they may be preceded by determiners. The second modification zone, called ‘descriptive’, encompasses the most central adjectives, that follow all the main criteria for adjectival status (they may occur in both the attributive and predicative position, can serve as conjoints in linked coordination, are gradable, etc.). The investigation discussed in the following sections of this article focuses on two types of adjectives belonging to the ‘descriptive zone’. Finally, ‘classifying’ adjectives ‘subcategorise

Noun Phrases with Prenominal and Postnominal Adjectives’, *Journal of Verbal Learning and Verbal Behavior*, 8 (1969), 471–80; James Martin and T. Ferb, ‘Contextual Factors in Preferred Adjective Ordering: A Critique’, *American Journal of Psychology*, 88 (1973), 201–15; R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik, *A Comprehensive Grammar of the English Language* (London: Longman, 1985); Meredith Richards, ‘The Pragmatic Rule of Adjective Ordering: A Critique’, *American Journal of Psychology*, 88 (1975), 201–15; R. Teschner and E. Evans, *Analyzing the Grammar of English* (Washington, D.C.: Georgetown University Press, 2007); Benjamin Whorf, ‘Grammatical Categories’, *Language*, 21 (1945), 1–11.

³ Hill and others, p. 192; Teschner and Evans, p. 147.

⁴ J. H. Danks and S. Glucksberg, ‘Psychological Scaling of Linguistic Properties’, *Language and Speech*, 13 (1970), 118–40.

⁵ Cf. Carl Bache, *The Order of Premodifying Adjectives in Present-day English* (Odense: Odense University Press, 1978); C. Bache and N. Davidsen-Neilsen, *Mastering English* (Berlin: Mouton de Gruyter, 1997); David Kemmerer, ‘Selective Impairment of Knowledge Underlying Prenominal Adjective Order: Evidence for the Autonomy of Grammatical Semantics’, *Journal of Neurolinguistics*, 13 (2000), 57–82; D. Kemmerer, C. Weber-Fox, K. Price, C. Zdanczyk, and H. Way, ‘Big Brown Dog or Brown Big Dog? An Electrophysiological Study of Semantic Constraints on Prenominal Adjective Order’, *Brain and Language*, 100 (2007), 238–56; D. Kemmerer, D. Tranel and C. Zdanczyk, ‘Knowledge of the Semantic Constraints on Adjective Order can be Selectively Impaired’, *Journal of Neurolinguistics*, 22 (2009), 91–108; Quirk and others.

⁶ Bache and Davidsen-Neilsen, p. 458.

the head they modify — e.g. “a medical dictionary” is a special kind of dictionary and “solar energy” is a special kind of energy. Classifying adjectives thus help establish precisely what sort of thing is involved in the expression’.⁷

In addition to the order resulting from the zone a given adjective belongs to, there are also rules that apply within zones. This is particularly noticeable among ‘descriptive’ adjectives. Still, Kemmerer provides empirical evidence that the level of inter-zone organisation has priority over intra-zone organisation.⁸ In his experiment on brain-damaged subjects mistakes were more frequent inside the ‘descriptive’ zone than between zones. This was also true for normal control subjects and it suggests that the two types of rules are stored in distinct neural networks and inter-zone distinctions are more recognisable than intra-zone distinctions. It is also worth adding that only some semantic features are visible to syntax, as was initially proposed by Pinker in his Grammatically Relevant Semantic Subsystems Hypothesis and later substantiated by Kemmerer in a series of publications through the first decade of this century.⁹ For instance, the fact that adjectives denoting age usually precede adjectives referring to colours proves that these semantic categories constrain linear order. On the other hand, the distinction between ‘red’ and ‘green’ is invisible to syntax and there are no preferences for one of them to follow the other. It is, therefore, apparent that many semantic differences do not influence linear sequence.

Table 1. Sequential orders proposed in selected publications. ‘-’ separates different slots in the sequence and ‘/’ denotes that no particular order for a given pair or group of adjectives was suggested.

Author(s)	Suggested sequential order of prenominal adjectives
Alexander ¹⁰	opinion – size – age – shape – colour – origin/past participle – noun
Campbell ¹¹	opinion – shape – age – colour – origin – material
Cinque ¹²	quantification – quality – size – shape – colour – nationality

⁷ Bache and others, p. 458.

⁸ Kemmerer, ‘Selective Impairment of Knowledge’.

⁹ Steven Pinker, *Learnability and Cognition* (Cambridge, MA: MIT Press, 1989); David Kemmerer, ‘Grammatically Relevant and Grammatically Irrelevant Features of Verb Meaning can be Independently Impaired’, *Aphasiology*, 14 (2000), 997–1020; ‘Selective Impairment of Knowledge’, ‘Neuropsychological Evidence for the Distinction between Grammatically Relevant and Irrelevant Components of Meaning’ (Commentary on R. Jackendoff ‘Precis of Foundations of Language’), *Behavioral and Brain Sciences*, 26 (2003), 684–85; ‘Why Can you Hit Someone on the Arm but not Break Someone on the Arm? A Neuropsychological Investigation of the English Body-Part Possessor Ascension Construction’, *Journal of Neurolinguistics*, 16 (2003), 13–36; ‘Action Verbs, Argument Structure Constructions, and the Mirror Neuron System’, in *From Action to Language via the Mirror Neuron System*, ed. by Michael Arbib (Cambridge: Cambridge University Press, 2006), pp. 347–73; D. Kemmerer and S. K. Wright, ‘Selective Impairment of Knowledge Underlying Un-prefixation: Further Evidence for the Autonomy of Grammatical Semantics’, *Journal of Neurolinguistics*, 15 (2002), 403–32; Kemmerer and others, ‘Big Brown Dog or Brown Big Dog?’; Kemmerer and others, ‘Knowledge of the Semantic Constraints on Adjective Order’.

¹⁰ *Longman English Grammar Practice*.

¹¹ Doug Campbell, *Professor Grammar’s Rule Book* ([London]: BBC, 1991)

¹² Guglielmo Cinque, ‘On the Evidence for Partial N-movement in the Romance DP’, in *Paths towards Universal*

Diachronic Development of the Order of Prenominal Adjectives in English

Author(s)	Suggested sequential order of prenominal adjectives
Hare and Wayne ¹³	size – colour – material
Dixon ¹⁴	value – dimension – physical property – speed – human propensity – age – colour
Evans ¹⁵	opinion – size/weight – age – shape – colour/temperature – participle – origin/nationality – material
Evans ¹⁶	opinion – size – age – shape – colour – origin – material – used for/be about (purpose)
Hewings ¹⁷	gradable – ungradable opinion – size/physical quality/shape/age – colour – participle – origin – material – type – purpose
Hill and others ¹⁸	general adjective – specific adjective
Hopkins and Hopkins ¹⁹	describers – classifiers describers: opinion – size – age – shape – colour classifiers: nationality – material – type
Kemmerer and others ²⁰	value – size – dimension – various physical properties – colour
Kingsbury and Wellman ²¹	subjective comment – size – age – shape – colour – nationality/origin – material
Lockhart and Martin ²²	high-ranking (less definite meaning) – low-ranking (more definite meaning)
Mańczak-Wohlfeld and others ²³	opinion – size – shape – age – colour – origin – substance – gerund
Parrott ²⁴	size – shape – colour – origin – material – use
Scheffehn ²⁵	size – colour – material (although 'colour – size – material' also encountered)

Grammar: Studies in Honor of Richard S. Kayne, ed. by G. Cinque and others (Washington, DC: Georgetown University Press, 1994), pp. 85–110.

¹³ V. C. Hare and O. Wayne, 'Development of Preferred Adjective Ordering in Children, Grades One to Five', *The Journal of Educational Research*, 71.4 (1978), 190–93.

¹⁴ Robert Dixon, *Where Have All the Adjectives Gone?* (Berlin: Mouton de Gruyter, 1982).

¹⁵ *CPE Use of English*.

¹⁶ *FCE Use of English*.

¹⁷ *Advanced Grammar in Use*.

¹⁸ *Grammar and Practice*.

¹⁹ *Developing Grammar in Context*.

²⁰ 'Big Brown Dog or Brown Big Dog?' and 'Knowledge of the Semantic Constraints'.

²¹ *Longman Advanced English*.

²² R. Lockhart and J. Martin, 'Adjective Order and the Recall of Adjective-Noun Triples', *Journal of Verbal Learning and Verbal Behavior*, 8 (1969), 272–75.

²³ *A Practical Grammar of English*.

²⁴ *Grammar for English Language Teachers*.

²⁵ Margaret Scheffelin, 'Children's Understanding on Constraints upon Adjective Order', *Journal of Learning Disabilities*, 4 (1971), 34–42.

Author(s)	Suggested sequential order of prenominal adjectives
Scott ²⁶	comment – size – length – height – speed – width – weight – temperature – age – shape – colour – nationality/origin – material
Scott-Barrett ²⁷	judgement – dimensions – colour – origin – material
Swan	age/shape/size/temperature – colour – origin – material – purpose
Yule ²⁸	describing – classifying describing: opinion – size – physical quality – age/time – shape – colour classifying: location – origin/source – material – type – purpose

2 Aims of the project

Among the linear orders summarised in Table 1 one may find various discrepancies. Firstly, different authors use different semantic categories. Secondly, there are instances in which the semantic categories are the same but labelled differently. Moreover, in some cases the patterns do not specify the preferred order for a given category, while in others such an order is explicitly proposed. Additionally, there are also instances in which selected semantic categories are placed in different slots in the sequential order. Of particular interest to the present study is the case of adjectives referring to the notions of AGE and SHAPE. Authors such as Alexander, Evans, Hopkins and Hopkins, Kingsbury and Wellman, Scott, and Yule suggest that the preferred order is AGE – SHAPE, while Campbell and Mańczak-Wohlfeld and others propose that the sequence SHAPE – AGE is more natural. Because of such discrepancies the initial aim of the research described in the following sections of this article is to establish the actual linear sequence of the adjectives denoting AGE and SHAPE in contemporary English. The second and primary aim of this publication concentrates on a more general problem: if both of the proposed orders appear in contemporary English (which is suggested by the results of the first part of the research described below), it may indicate a diachronic change. Therefore, it is necessary also to investigate possible trends in the historical development of the two adjective orders. Any consistent patterns observed in such a diachronic investigation may contribute to our understanding of the way languages change over time.

3 Methods

In order to accomplish the aims of this project, the preferred order of adjective types under discussion must be tested in a large corpus of English which allows diachronic analysis. One

²⁶ Gary-John Scott, 'Stacked Adjectival Modification and the Structure of Nominal Phrases', in *Functional Structure in DP and IP*, ed. by Cinque Guglielmo (New York: Oxford University Press, 2002), pp. 91–120.

²⁷ *New Proficiency Use of English*.

²⁸ *Oxford Practice Grammar*.

²⁹ Mark Davies, *The Corpus of Historical American English: 400 Million Words, 1810–2009* (2010–),

possibility would be to use *COHA (The Corpus of Historical American English)*.²⁹ It is based on 400 million words from 1810 to 2009. Even though the corpus seems large, however, pilot studies revealed that it is still too small of statistically sound evaluation of both the synchronic distribution and the historical development of the two orders of adjectives. Therefore, the American version of the *Google Books Corpus*, based on the *Google Books Ngram Corpus*, was chosen instead.³⁰

The first edition of the *Google Books Ngram Corpus* has been available online since 2010.³¹ The first suggested application was the study of culture and several papers investigating such aspects with the use of *Google Books Ngrams* have recently been published.³² J. B. Michel and others have also given examples of the ways in which the corpus may be utilised in lexicography and in the analysis of the evolution of grammar.³³ Since their publication, numerous studies focusing on the use of *Google Books Ngrams* for linguistic purposes have been conducted. This includes disciplines such as syntax,³⁴ semantics,³⁵ and psycholinguistics.³⁶ Moreover, the Ngrams have been used in several studies on the diachronic development of English. For instance, Demetris Koutsoyiannis analysed the historical development of several expressions with the word 'change',³⁷ while R. Mihalcea and C. Nastase have conducted a diachronic study of 200 words chosen to represent the four basic parts of speech: nouns, verbs, adjectives and adverbs.³⁸ Furthermore, K. Gulordava and M. Baroni proposed an automatic detection system for semantic change.³⁹ They applied the

<http://corpus.byu.edu/coha/> [accessed 15 July 2013].

³⁰ Mark Davies, *Google Books Corpus, Based on Google Books N-grams* (2011–), <http://googlebooks.byu.edu> [accessed 15 July 2013].

³¹ Its structure and basic use is described by J. B. Michel and others, 'Quantitative Analysis of Culture Using Millions of Digitized Books', *Science*, 331 [6014] (2011), 176–82.

³² Patrick Juola, 'Using the Google N-Gram Corpus to Measure Cultural Complexity', *Literary and Linguistic Computing*, 28 (2013), 668–75; Martin Ravallion, 'The Two Poverty Enlightenments', *The World Bank Policy Research Working Paper*, 5549 (2011); D. S. Soper and O. Turel, 'An N-gram Analysis of Communications 2000–2010', *Communications of the ACM*, 55.5 (2012), 81–87.

³³ 'Quantitative Analysis of Culture'.

³⁴ Y. Goldberg and J. Orwant, 'A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books', in *Proceedings of the Second Joint Conference on Lexical and Computational Semantics*, 2 vols (Stroudsburg, PA: Association for Computational Linguistics, 2013), I 241–47.

³⁵ K. Gulordava and M. Baroni, 'A Distributional Similarity Approach to the Detection of Semantic Change in the Google Books Ngram Corpus', in *Proceedings of the GEMS 2011 Workshop on Geometrical Models of Natural Language Semantics* (Stroudsburg, PA: Association for Computational Linguistics, 2011), pp. 67–71; C. Joubarne and D. Inkpen, 'Comparison of Semantic Similarity for Different Languages Using the Google N-gram Corpus and Second-Order Co-occurrence Measures', in *Proceedings of the 24th Canadian Conference on Advances in Artificial Intelligence* (Berlin: Springer, 2011), pp. 216–21; A. Islam, E. Milios and V. Kešelj, 'Comparing Word Relatedness Measures Based on Google N-grams', in *Proceedings of COLING 2012 International Conference on Computational Linguistics: Posters* (Mumbai: The COLING 2012 Organizing Committee, 2012), pp. 495–506; H. Agt and R. D. Kutsche, 'Automated Construction of a Large Semantic Network of Related Terms for Domain-Specific Modeling', in *Advanced Information Systems Engineering* (Berlin: Springer, 2013), pp. 610–25.

³⁶ M. Brysbaert, E. Keuleers and B. New, 'Assessing the Usefulness of Google Books' Word Frequencies for Psycholinguistic Research on Word Processing', *Frontiers in Psychology*, 2 (2011), <https://doi.org/10.3389/fpsyg.2011.00027>; E. Keuleers, M. Brysbaert and B. New, 'An Evaluation of the Google Books Ngrams for Psycholinguistic Research', *Lexical Resources in Psycholinguistic Research*, 3 (2011), 23–26.

³⁷ Demetris Koutsoyiannis, 'Hydrology and Change', *Hydrological Sciences Journal*, 58 (2013), 1177–97.

³⁸ R. Mihalcea and V. Nastase, 'Word Epoch Disambiguation: Finding How Words Change over Time', in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, 2 vols (Stroudsburg, PA: Association for Computational Linguistics, 2012), II 259–63.

³⁹ 'A Distributional Similarity Approach'.

distributional semantics model suggested by P. D. Turney and P. Pantel and used the portion of the corpus that includes American English 2-grams.⁴⁰

Since its initial launch in 2010, *Google Books Ngrams* have undergone numerous changes. Y. Lin and others describe an improved edition of the corpus which is larger and includes tagging of basic parts of speech as well as syntactic annotations.⁴¹ The Ngrams have also been integrated into the interface created by Mark Davies.⁴² This version of the corpus, which in this paper is referred to as the *Google Books Corpus*, makes it possible to search the data in a more flexible manner because it uses the same advanced part-of-speech tagging system as Davies's aforementioned *Corpus of Historical American English* and his *Corpus of Contemporary American English*.⁴³

The investigation described in Section 4 was conducted on the *Google Books Corpus*, but it must be stressed that even the advanced interface created by Mark Davies does not include semantic annotation, so it was not possible, for example, to search for 'any adjective' referring to AGE followed by 'any adjective' denoting SHAPE. Instead, selected pairs of adjectives needed to be typed in manually. As a consequence, a list of such possible pairs was compiled. It included five adjectives denoting AGE (OR TIME — the distinction between these two semantic categories is not made in the literature): 'new', 'ancient', 'modern', 'current' and 'contemporary'. Many other adjectives were rejected for various reasons. Some of them were likely to have two (or more) very dissimilar meanings, as in the case of the word 'old', which basically refers to any entity which has existed for many years, but may also be used in expressions such as 'an old friend' with the meaning 'very familiar'. Another problem concerned the fact that many adjectives referring to AGE (OR TIME) are also used as other parts of speech. For instance, 'present' is an adjective in 'a present moment', a noun in 'a birthday present' or a verb in 'to present the results'. Such problems could theoretically be solved by a detailed qualitative analysis of individual examples, in which case adjectives such as 'old' or 'present' could still be included in the investigation. However, for technical reasons, such a solution proved impossible. The *Google Books Corpus* is based on only a part of the *Google Books* depository; the latter is constantly changing and new items are being continually added to it. This means that any query run from the *Google Books Corpus* search engine will give a potentially smaller number of tokens than the ones found in the current *Google Books* depository. As a result, individual examples of combinations of adjectives such as 'old' and 'present' with other adjectives found in the *Google Books Corpus* cannot be effectively evaluated qualitatively. There is also a mismatch between the results obtained in the *Google Books Corpus* and the *Google Books* site resulting from the fact that the former uses only one set of data (American English in our case), while the latter provides extracts from all available books it has access to. As a consequence of all these inconsistencies, a systematic, qualitative analysis was unfeasible and the final decision was made to choose only the adjectives which referred to the notion of AGE in the least ambiguous way. Still, it must be stressed that even a very careful selection of examples does not guarantee the appropriateness of the tokens found

⁴⁰ P. D. Turney and P. Pantel, 'From Frequency to Meaning: Vector Space Models of Semantics', *Journal of Artificial Intelligence Research*, 37 (2010), 141–88.

⁴¹ Y. Lin and others, 'Syntactic Annotations for the Google Books Ngram Corpus' in *Proceedings of the ACL 2012 System Demonstrations* (Stroudsburg, PA: Association for Computational Linguistics, 2012), pp. 169–74.

⁴² *Google Books Corpus*.

⁴³ Mark Davies, *The Corpus of Contemporary American English: 520 Million words, 1990–present* (2008–), available online at <http://corpus.byu.edu/coca/> [accessed 15 July 2013].

in the corpus. Other unpredictable factors may influence the results. For example, it is difficult to decide whether the adjective 'modern' could not in some cases be interpreted as evaluative, since in some contexts 'modernity' is regarded as something positive in contrast to anything which is 'old-fashioned'. For such reasons, a potential degree of error of the results must be presumed.

As far as the semantic category of *SHAPE* is concerned, some examples were excluded from the analysis for reasons similar to those described in the previous paragraph (e.g. the words 'square' and 'oblong' may function as nouns as well as adjectives), and, eventually, the following seven were chosen: 'rectangular', 'x-shaped', 'jagged', 'triangular', 'curved', 'conical' and 'spherical'. According to the online search engine *wordandphrase.info*, which provides information on the relative frequencies of words in English based on data from the *Corpus of Contemporary American English*, these were the most frequent adjectives from among the possible candidates.⁴⁴ Other possible adjectives, such as 'ovate', were too rare for the purposes of the current project.

All of the selected adjectives denoting *AGE* were juxtaposed with all of the adjectives referring to *SHAPE* in both of the tested word orders, resulting in 70 searches (35 for each order). Nevertheless, only some of the combinations used in the investigation were actually present in the corpus (see Section 4).

The combinations were typed in the main search box of the corpus as pairs of words without any additional annotation. The only exception was the group of adjectives of the 'x-shaped' type which involve any lexical element preceding 'shaped' (e.g. 'heart-shaped', 'diamond-shaped', etc.). In order to obtain all possible combinations of the *AGE* adjectives with such compounds, the symbol used in the investigation was '*-shaped', which is interpreted by the corpus search engine as 'any word' before '-shaped'.

Finally, it should be pointed out that for the purposes of calculating the percentages from the empirical part of the paper, the corpus samples were divided by 2. For instance, for the most recent decade in the American version of the Google Books Corpus, which was used in Section 4.1, the percentages were calculated out of 13441150000, rather than 26882300000 tokens. This solution does not change the relative differences between individual results, but it provides a more realistic picture of frequencies, since the items under analysis are pairs of words, not single words.

4 Results and Discussion

4.1 Synchronic analysis

As stated in Section 2, before conducting a diachronic investigation on the development of the two adjective orders under discussion, an analysis of the synchronic distribution of the two proposals in contemporary English was performed. The results of this brief investigation are shown in Table 2.

It is immediately visible that the order *AGE* – *SHAPE* is more frequent than *SHAPE* – *AGE*. While the former was found in as many as 12 out of the 70 different combinations investigated, the latter was encountered in only 1 combination. Moreover, 437 tokens exhibited the prevailing order, which is almost 100 times more frequent than the 5 cases

⁴⁴ <http://www.wordandphrase.info/frequencyList.asp>.

Order AGE – SHAPE		Order SHAPE – AGE	
new rectangular	120	rectangular new	5
new { <i>x</i> }-shaped	14		
new triangular	103		
new curved	93		
new conical	19		
new spherical	26		
ancient rectangular	11		
ancient triangular	6		
modern rectangular	11		
modern curved	15		
modern conical	5		
current rectangular	11		
sum	437		

Table 2. The two adjective orders in contemporary American English

of the SHAPE – AGE pattern. In terms of percentages, the two results may be expressed as 0.000003258649744% and 0.000000037199198%, respectively. The 95% confidence interval for the difference between them is 0.000003221450546% \pm 0.000000306917394%, and the p-value for this difference is smaller than 0.0001. Therefore, the predominance of the AGE – SHAPE adjective order in contemporary American English is beyond any doubt.

It is worth noting at this point that the number of tokens for examples with ‘new’ is relatively higher than for other combinations encountered in the *Google Books Corpus* and constitutes around 80% of all the tokens listed in Table 2. One possible explanation for such a popularity of combinations with ‘new’ is that the adjective is generally very common in English. The list of the top 5000 words available at *wordfrequency.info*, which is based on the *Corpus of Contemporary American English*, indicates that ‘new’ is the second most frequent adjective after ‘other’.

It must be added that the small number of SHAPE – AGE tokens does not necessarily mean that they are that infrequent on the *Google Books* site. One of the restrictions imposed on the *Google Books Corpus* is that it is based on the n-grams which occur at least 40 times or more. They are, in fact, the same n-grams which are available on the *Google Books N-gram Viewer*. Any number of tokens below this threshold is not displayed in the *Google Books Corpus* (the result for ‘rectangular new’ was 5, but it was established on the basis of all the examples for the last two centuries, and their total number exceeded 40), which may cause bias in the results, especially as regards pairs that represent rare phenomena in the English language. Consequently, the order SHAPE – AGE may be used more often in contemporary English than the data in Table 2 suggest.

4.2 Diachronic analysis

Table 3 summarises the raw occurrences of the two adjective orders in the American portion of the *Google Books Corpus* over the last two hundred years. As in the synchronic analysis discussed in Section 4.1, the pattern SHAPE – AGE was found only for the example ‘rectangular

Diachronic Development of the Order of Prenominal Adjectives in English

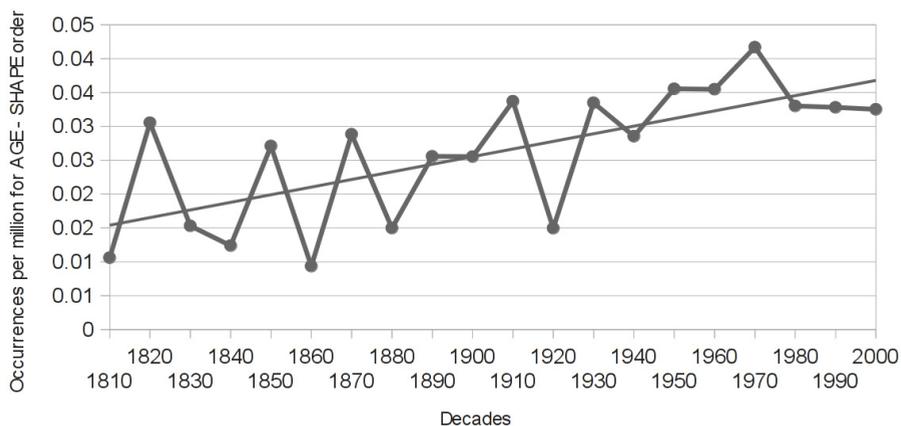


Figure 1. Normalised frequencies of the pairs of adjectives following the order AGE – SHAPE. The slanting line in the middle represents linear regression.

new’. None of the other pairs was encountered in this order, but, similarly to the problem encountered in the previous section, some of the tokens were inevitably missed due to the threshold limitation of 40 tokens. Nevertheless, the preference for the order of AGE – SHAPE is beyond any reasonable doubt.

Among the pairs of adjectives representing the AGE – SHAPE pattern, combinations involving ‘new’ were the most frequent. Out of the 9 different combinations, as many as 7 contained this adjective. This distribution is the result of the general high frequency of the adjective “new” mentioned in the previous section.

The data in Table 3 present raw frequency counts and they must be normalised before any valid conclusions may be drawn. This has been done in Figure 1, which summarises the relative occurrences of the pairs of adjectives representing the AGE – SHAPE pattern in the decades from 1810 to 2000. Although the line showing the frequency of use of this adjective order meanders, the graph suggests an overall upward trend. This may be tested statistically in at least two different ways. Firstly, it is possible to calculate the percentages of occurrence for all the tokens for the nineteenth century (0.00000207985395%, or 0.0207985395 per million) and the tokens for the twentieth century (0.00000335089772%, or 0.0335089772 per million). It is plainly visible that the AGE – SHAPE adjective order was less common in the former than in the latter. The 95% confidence interval for this difference is 0.00000127104377% ± 0.00000007563067995%, and the p-value is lower than 0.0001, which leaves no doubt as to the statistical significance of the difference between the two results. The AGE – SHAPE pattern was less popular in the nineteenth century than in the twentieth. Secondly, the chi-squared test for trend in proportions also indicates that there is a statistically relevant change: the obtained p-value is smaller than 0.0001.

Due to the limited number of examples, the opposite tendency for the SHAPE – AGE order is more difficult to ascertain. While the distribution of the 42 examples of ‘rectangular new’

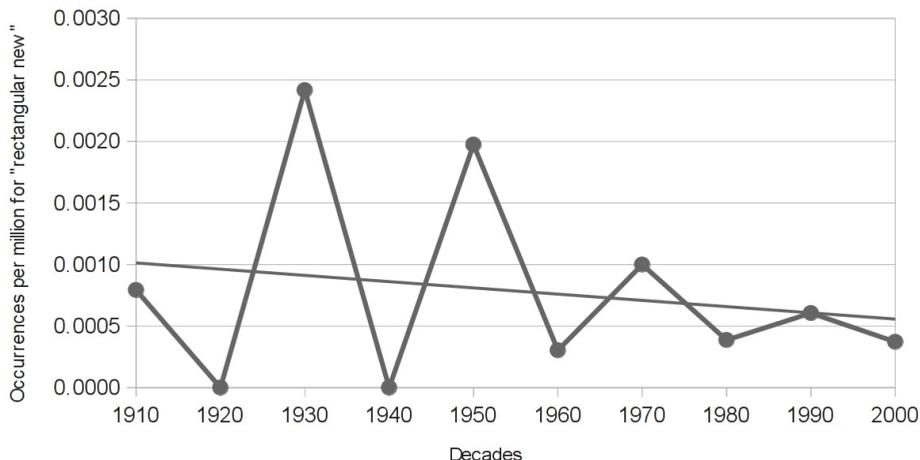


Figure 2. Normalised frequencies of 'rectangular new'. The slanting line in the middle represents linear regression.

in the twentieth century indicate a downward trend (see Figure 2), the chi-squared test for trend in proportions does not yield a statistically significant result in this case ($p = 0.1145$). The limitation of 40 tokens imposed on *the Google Books Corpus* resulted in a small number of examples of the *SHAPE – AGE* pattern and the diachronic development of this word order could not be adequately investigated.

On balance, the results of the present investigation confirm an increase in the use of the *AGE – SHAPE* adjective pattern over the last two centuries. It must be emphasised, though, that this claim does not imply that the pairs of adjectives of the *AGE – SHAPE* type are becoming more popular in absolute terms. Even though the opposite trend to use the *SHAPE – AGE* order less frequently could not be proven statistically, the assumption that it is happening is more reasonable than concluding that American writers have been mentioning more and more objects which are, for instance, new and triangular. This would suggest that there are proportionally more new triangular objects in existence now than a hundred years ago, or that such objects are noticed more often. With the methodology applied in the present study, this interpretation is theoretically also possible and cannot be completely excluded; however, Figures 3 and 4 provide additional evidence against such an explanation. Figure 3 presents the diachronic development of nine pairs of adjectives of the *AGE – SHAPE* type. All of them reveal an upward trend, although it is easily noticeable only for 'new rectangular', 'new triangular' and 'new curved'. The results of the chi-squared test for trend in proportions confirm these observations. The p -values for all nine cases are below 0.0001. The pair 'new conical' is not included in Figure 3 because its diachronic development does not follow any consistent pattern, but the three remaining pairs shown in Figure 4 seem to exhibit a downward trend. It is especially evident in the case of the pair 'new spherical'. The chi-squared test for trend in proportions confirms this observation ($p < 0.0001$), but in the other two examples, the weak

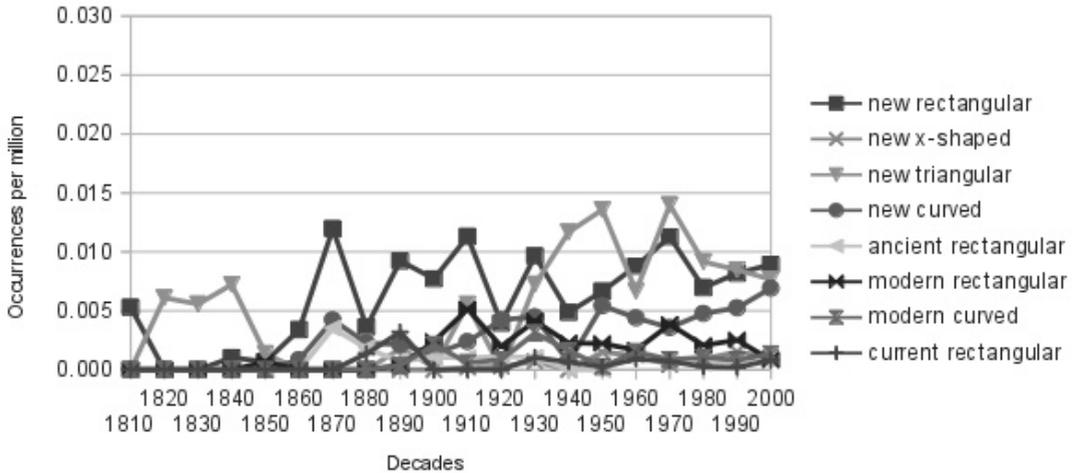


Figure 3. Normalised frequencies of individual adjective pairs exhibiting an upward trend.

negative trend cannot be statistically validated. The p-values for both ‘ancient triangular’ and ‘modern conical’ are clearly above the alpha level of 0.05 (0.1067 and 0.2903, respectively). This leads to the conclusion that the general tendency to use the AGE – SHAPE pattern applies to the majority of the adjective pairs investigated in the current study and not just selected examples. The alternative ‘absolute interpretation’, according to which there are more objects with the qualities indicated by the adjective pairs in existence now than a hundred years ago, would have to assume an increase in the number of objects which are ‘new rectangular’, ‘new x-shaped’, ‘new triangular’, ‘new curved’, ‘new oblong’, ‘ancient rectangular’, ‘modern rectangular’, ‘modern curved’ and ‘current rectangular’, and explain why only the objects which are ‘new spherical’ have decreased in popularity. The ultimate interpretation proposed in this paper is, therefore, less controversial. It is more reasonable to assume that we are dealing with a change in the order of adjectives rather than an increase in frequencies of individual word pairs.

Conclusion

The present study has found that the adjective order AGE – SHAPE is prevalent in contemporary American English. This agrees with most of the proposals put forward by the authors discussed in Section 2. The only exceptions were the suggestions expressed by Campbell and Mańczak-Wohlfeld and others,⁴⁵ according to whom the preferred order should be SHAPE – AGE. The question which arises is why these two authors’ position was different from those expressed in other publications. Firstly, it must be stressed that the SHAPE – AGE order is also found in contemporary English. The claim that it should be used is, therefore, not entirely

⁴⁵ Campbell, *Professor Grammar’s Rule Book*; Mańczak-Wohlfeld and others, *A Practical Grammar of English*.

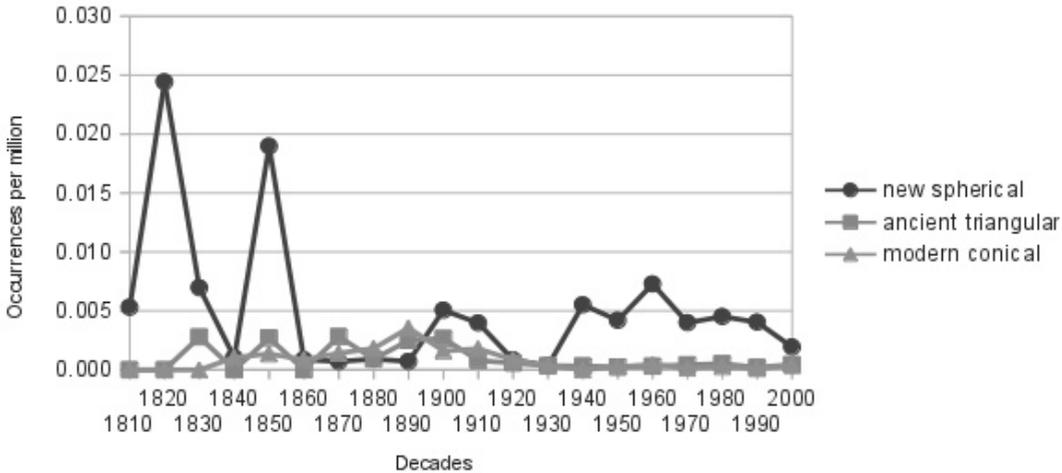


Figure 4. Normalised frequencies of individual adjective pairs exhibiting a downward trend.

unsubstantiated. The problem is that this pattern is clearly less popular than *AGE – SHAPE* and advising use of the opposite is, ultimately, wrong. A possible reason for such a mistake could be the fact that both Campbell and Mańczak-Wohlfeld and others did not rely on empirical data. At the time of their publications, the use of electronic corpora was severely limited and the corpora which were available were significantly smaller than the ones used in this project. Therefore, claims made at the time about a given grammatical rule tended to rely on author’s intuition. Thus, these results also point to the superiority of empirical methods over introspection.

It has also been shown that the *AGE – SHAPE* pattern has increased in popularity over the nineteenth and twentieth centuries. Such an observation enhances our understanding of language change. The preferred orders of adjectives in the attributive position may alter in time just as is the case with many other aspects of human communication. While some facets of the phenomenon under discussion tend to be universal across languages (cf. Svenonius 2008), others may be more language specific and change in the course of diachronic development.

The reverse order *SHAPE – AGE* was only observed a few times. As a result, its historical decline was impossible to prove statistically with the methodology applied above, but the additional aspects discussed in Section 4.2 support such an interpretation.

One of the possible aims for the future is to investigate this issue with the use of methods which would not be hampered by the restrictions described in Section 4.1. Such a solution requires new advances in corpus linguistics resources. Furthermore, it would be interesting to study the diachronic development of selected adjective orders in dialects of English other than American. This could be partially accomplished with the use of some currently available corpora, but, again, in order to obtain statistically reliable results on both adjective orders under discussion, larger corpora are required.

Table 3. Raw frequencies of the pairs of adjectives under analysis in the *Google Books Corpus*

AGE – SHAPE	Total	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
new rectangular	620	1			1	1	4	17	8	26	29	57	14	28	15	27	58	79	54	81	120
new X-shaped	57											1	1	2		6	8	3	7	15	14
new triangular	561		2	4	7	2				5	1	28		21	36	55	44	98	71	84	103
new curved	323						1	6	5	6	5	12	15	13	2	22	29	25	37	52	93
new oblong	59									3	1		5	1	6	2	7	2	9	22	1
new conical	104				2	2	4	6	5	3	5	5	1	3	3	4	7	12	16	7	19
new spherical	303	1	8	5	1	28	1	1	2	2	19	20	3	1	17	17	48	28	35	40	26
ancient rectangular	63							5	4	2	4	5	4	3			8	5	4	8	11
ancient triangular	55			2		4		4	2	7	10	4	2	1	1	1	2	3	4	2	6
modern rectangular	162					1				1	9	26	7	12	7	9	11	27	16	25	11
modern curved	80									1	8	3	3	9	5	1	10	6	8	8	18
modern conical	52				1	2	1	2	4	10	6	9	3	1		1	3	1	2	1	5
current rectangular	44								3	9				3	2	1	6	5	2	2	11
SHAPE – AGE																					
rectangular new	42											4		7		8	2	7	3	6	5