

## Synonym selection as a strategy of stress clash avoidance

Julia Schlüter & Gabriele Knappe  
Otto-Friedrich-Universität Bamberg

Published in:

Sebastian Hoffmann, Andrea Sand, Sabine Arndt-Lappe & Lisa Marie Dillmann (eds.)  
2018. *Corpora and Lexis*. Leiden & Boston: Brill/Rodopi. 69–105.

DOI: [https://doi.org/10.1163/9789004361133\\_005](https://doi.org/10.1163/9789004361133_005)

[NB: The paragraphs **Abstract** and **Keywords** are not published in the printed version.]

### Abstract

In this paper we argue that (near) synonymy can result in syntactic specialization, and we test the hypothesis that one major factor underlying such fixation may be rhythmic in nature. Thus, the analysis extends the study of the preference for alternating stressed and unstressed syllables (Schlüter, 2005) to lexical choices. We investigate the syntactic distribution of the (near-)synonymous adjective pairs *rich vs. wealthy*, *fast/quick vs. rapid*, *glad vs. happy* and *shut vs. closed*, showing that the monosyllabic members – or the one with a short syllable as opposed to a long syllable – tend to be underrepresented in prenominal position. Our hypothesis is that, due to the pervasiveness of initial stress in English nouns, monosyllabic adjectives tend to be avoided in prenominal position if they can be replaced by disyllabic equivalents. By extension, in the absence of such equivalents, other kinds of temporal spacing between stresses, such as long instead of short syllables, can satisfy the rhythmic requirement. Drawing on large historical as well as present-day corpora covering the 19<sup>th</sup> and 20<sup>th</sup> centuries (mainly COHA and COCA), we portray some relatively stable asymmetrical diachronic and synchronic distributions of these synonym pairs and triplets across attributive and predicative uses. The analysis will also point to potentially interacting factors, such as concomitant semantic specialization resulting from this syntactic fixation and the status of (fixed) collocations. The latter will be analysed in more diachronic depth on the basis of EEBO.

**Keywords:** Principle of Rhythmic Alternation, stress clashes, syllable duration, interaction across linguistic levels, lexical choices, synonyms, collocations, fixed expressions, phraseological units

### 1. Introduction and hypothesis

The Principle of Rhythmic Alternation holds that

[...] stressed and stressless syllables tend to alternate at *rhythmically ideal disyllabic distances*. Rhythmic alternation manifests itself by the avoidance of sequences of stressed syllables, as well as of long sequences of stressless syllables. (Kager, 1989, p. 2, italics in the original)<sup>1</sup>

Schlüter (2005) and subsequent studies<sup>2</sup> have provided quantitative evidence of the influence of rhythm on grammatical variation and change. This paper is an attempt to

<sup>1</sup> For some further, largely congruent, formulations of the principle, see Jespersen (1909, p. 156; 1972, p. 97); Fijn van Draat (1910, p. 9); Behaghel (1924, p. vi-vii); Bolinger (1965b, p. 139); Selkirk (1984, p. 37); Nespor & Vogel (1989, p. 69, 82); Couper-Kuhlen (1986, p. 60).

<sup>2</sup> Cf. Lee & Gibbons (2007); Schlüter (2008, 2015); Anttila, Adams & Speriosu (2010); Speyer (2010); Mollin

determine to what degree stress clash avoidance can also determine lexical choices, which is an area that has so far been almost completely neglected. The basic idea for this study is inspired by Bolinger's (1965b, p. 149) observation that *a glád dáy* seems objectionable, while *a glád occásion* or *a háppy dáy* are acceptable.<sup>3</sup> He surmises that in cases where speakers or writers plan to use a sequence of lexemes that would result in the adjacency of two strongly stressed syllables, they may replace one of the words with a rhythmically more fitting one. Needless to say, such a bypass will be adopted only if it does not substantially alter the intended meaning.

Putting Bolinger's impressionistic (albeit very perceptive) observation to an empirical test is not an easy task since in the lexical domain, semantic equivalence is harder to establish than in the domains of morphological and syntactic variants. To counterbalance these difficulties, we will focus on rhythmic and syntactic constellations that can be predicted to exhibit the strongest impact of the Principle of Rhythmic Alternation. Its effect strength has been shown to depend both on the degree of stress (primary, secondary, no stress) and the degree of syntactic bounding (i.e., the width of the prosodic juncture). We will adopt a binary division between primary and non-primary stress and focus on attributive adjectives immediately preceding the nouns they modify. Prenominal uses (also referred to as attributive uses in this paper) represent the critical contexts, where the two elements are syntactically and prosodically tightly bound to each other. Combinations of adjectives + nouns are relatively frequent, as c. 65 % of adjectives occur prenominal (according to a POS-tag-based search of COCA, on which see below). What makes the situation even more liable to stress clashes is that initial stress is prevalent in English nouns: Approximately 85 % of noun tokens in the Brown corpus have initial stress (cf. Schlüter, 2005, p. 330; cf. also Cutler & Carter, 1987). Two examples of such critical contexts are given in (1a) and (2a).

- |                           |                          |
|---------------------------|--------------------------|
| (1) a. <i>glád</i> péople | <i>háppy</i> péople      |
| b. <i>they were glád</i>  | <i>they were háppy</i>   |
| (2) a. <i>rích</i> péople | <i>wéalthy</i> péople    |
| b. <i>they were rích</i>  | <i>they were wéalthy</i> |

When adjectives appear in other, i.e. predicative and otherwise postnominal, positions, as in (1b) and (2b), they are typically followed by (unstressed) function words or pauses. Even in the few occurrences where this is not the case, they are rarely part of a phrase with the same degree of syntactic and prosodic bounding as an attributive structure. Thus, other contexts can be considered as rhythmically uncritical and will serve as control contexts. To test Bolinger's (1965b, p. 149) suggestion, we can therefore formulate two hypotheses, aiming to reject the null hypothesis and to support the alternative hypothesis:

- H<sub>0</sub>: Two adjectives with equivalent meanings and different rhythmic shapes (monosyllabic and end-stressed or plurisyllabic and non-end-stressed) occur equally often in all syntactic functions.
- H<sub>1</sub>: Two adjectives with equivalent meanings and different rhythmic shapes (monosyllabic and end-stressed or plurisyllabic and non-end-stressed) do not occur equally often in all syntactic functions.

---

(2012, 2014); Lohmann (2014); Ehret, Wolk & Szmrecsanyi (2014); Shih, Grafmiller, Futrell & Bresnan (2015).

<sup>3</sup> In the examples here and elsewhere, acute accents are added to indicate primary lexical stress.

If it was only the meaning of an adjective that determined its placement in sentences, details of its phonological (or rhythmic) shape should be irrelevant. But if H<sub>1</sub> holds true, we expect a monosyllabic or end-stressed adjective (like *glad* or *rich*) to occur less often in a rhythmically critical context than a non-end-stressed adjective with an equivalent meaning (like *happy* or *wealthy*). Their occurrence in control contexts should remain unaffected by their rhythmic shapes.

One complication of our proposal lies in the fact that words cannot be studied irrespective of their individual histories, including their baseline frequencies, deep-rooted co-occurrence patterns, established semantic ranges, etc. In order to assess the time-depth and relative stability of our findings for Present-Day English, we will therefore compare the syntactic uses of adjective pairs (and a triplet) over the past 200 years. In addition, we will, in a case study, fathom the role of specific fixed collocations (phraseological units) by looking at their developments from the 16<sup>th</sup> century onwards. If the Principle of Rhythmic Alternation turns out to have explanatory potential, we predict a contrast between critical and control contexts that should be stable diachronically, across different varieties (British and American English) and across spoken and written modes, as discussed in the concluding section.

Our prediction is reflected in the corpus selection and search procedures described in section 3. Before turning to the considerable methodological challenges posed by our research question, in the following section 2, we will briefly discuss our choice of adjectives and the notion of synonymy that we employ. Section 4 looks at the results for the synonym pairs/triplets individually and for *happy* – *glad* in greater historical depth, and ends with a synopsis of spoken and written data from Present-Day English. In the concluding section, we will throw into relief the theoretical relevance of our findings and point out shortcomings and avenues for further research.

## 2 Choice of items for analysis and notes on synonymy

Besides the seminal remarks in Bolinger (1965b, p. 149), who – among some other examples – draws attention to the pair *glad* – *happy*, the small selection of adjectives to be scrutinized in this paper takes *Oxford Learner's Dictionaries Online* as its starting point. This resource contains a list of lemmas where special notes about synonyms, defined as “words with similar meanings”,<sup>4</sup> can be found. From the considerable number of adjectives listed, we selected the pair *rich* – *wealthy* and the triplet *fast* – *quick* – *rapid* on the grounds that the adjectives a) involve a rhythmically relevant difference, b) constitute sufficiently close synonyms, and c) can be retrieved from a corpus with reasonable precision. (For some additional candidates, see section 5.) The pair *shut* – *closed*, which will receive special consideration in section 4.4, is adduced by Bolinger (1965a, p. 135) in a different but comparable context: Though both members of the pair are monosyllables, they will allow us to extend and supplement our notion of rhythmic alternation to include the impact of syllable length.

Following Cruse (2002, p. 486), we understand synonymy, as “a relation of similarity/identity of meaning between senses associated with two (or more) lexical forms”, or, in cognitive terms, two or more lexical forms that are mapped onto the same concept. For our analysis of pairs (and the triplet) of adjectives, Cruse’s (1986, p. 88) notion of cognitive synonymy is particularly useful (cf. also Paradis, 1997, p. 67; for further discussion cf. Cruse, 2002, p. 489f. on propositional synonymy):

---

<sup>4</sup> Cf. the “usage notes” at: [www.oxfordlearnersdictionaries.com/wordlist/english/usage\\_notes/UsageNoteList\\_Synonyms/](http://www.oxfordlearnersdictionaries.com/wordlist/english/usage_notes/UsageNoteList_Synonyms/).

X is a cognitive synonym of Y if (i) X and Y are syntactically identical, and (ii) any grammatical declarative sentence S containing X has equivalent truth-conditions to another sentence S<sup>1</sup>, which is identical to S except that X is replaced by Y.

The adjectives we consider may be interchanged in declarative sentences without changing the equivalence of the truth conditions of the sentences involved, and this goes for both types of syntactic contexts (i.e., prenominal and postnominal). The following examples will show the relevance of Cruse's cognitive synonymy in terms of mutual entailment for our approach:<sup>5</sup>

- (3) a. The Arabian sheik was **wealthy**.  
 b. The Arabian sheik was **rich**.  
 (4) a. He was a **wealthy** man.  
 b. He was a **rich** man.

If sentence (3a) is true, sentence (3b) is true, and if sentence (3b) is true, sentence (3a) is true; if sentence (3a) is false, sentence (3b) is false, and if sentence (3b) is false, sentence (3a) is false. The same mutual entailment can be shown for (4a) and (4b). Since all conditions apply, *wealthy* and *rich* are cognitive synonyms in the contexts relevant for our analysis. As opposed to absolute synonymy, not all senses of the lexemes have to match in this way; the denotational ranges of the lexemes may in fact differ in size (cf. also section 4.1). In addition, the notion of cognitive synonymy tolerates collocational restrictions (e.g., *filthy rich*, but \**filthy wealthy*; cf. Cruse, 1986, p. 279-282) and will for the present purposes be extended to cover the merely partial equivalence of prenominal *happy* and *glad* (cf. 4.3).

Cognitive synonyms form the second of three levels of synonyms in Cruse, which are differentiated by frequency. The most frequent kind of synonymy is plesionymy (near synonymy; cf. Cruse, 2002, p. 490f.), where the senses connected to word-forms show a relatively minor meaning difference. The relation of plesionymy can be found out by the *or rather* test, as in *He was murdered, or rather executed*. Plesionyms yield sentences with different truth-conditions and are therefore not relevant for our study. The third type, absolute synonyms – “two distinct lexical forms being associated with senses that are identical” (Cruse, 2002, p. 488) – belong into the smallest, “vanishingly rare” (Cruse 2002, p. 490) group of synonyms.

“Absolute synonymy is rare – and when found mostly fleeting” (Dolezal, 2013, p. 255). This transience as well as instances of cognitive synonymy may give way to splits. Two types of splits have been discussed in the literature so far: First, they can be of a semantic nature (cf. Cruse, 2002, p. 489), and possibly result in stylistic divergence. For instance, the native word *hearty* and the foreign loan *cordial* both were in late Middle English used to express feelings that were supposed to spring from the heart. But then *hearty* came to imply a certain physical vigour and downrightness, as in *a hearty dinner*, and *cordial* a more quiet or conventional manifestation, as in *a cordial reception* (cf. Baugh & Cable, 2013, p. 176; *cordial* is labelled “formal” in *Oxford Learner's Dictionaries Online*). A second form the split can take is (idiosyncratic) collocational restriction, for example in the pair *flawless* and *impeccable*. *Flawless*, according to Cruse (2002, p. 489), shows collocational restrictions: While the noun *performance* can be modified by either

---

<sup>5</sup> Sentences (3a) and (3b) were used by Lyons (1968, p. 450) to illustrate his concept of synonymy, which is propositional synonymy.

adjective, *impeccable*, but not *flawless*, collocates with the noun *behaviour*.<sup>6</sup> The present contribution suggests a third type of split, namely syntactic specialization, and investigates the factors underlying such instances of syntactic split.

### 3. Corpora, search expressions and sampling methods

In order to test the assumed diachronic stability of the contrasting distributions of synonymous adjectives, the main parts of the following analyses (4.1-4.4) are based on a diachronic corpus spanning two centuries: The *Corpus of Historical American English* (COHA) comprises a total of 400 million words of written American English taken from fiction, popular magazines, newspapers and non-fiction books, and is subdivided into 20 decades (1810s-2000s). To manage the large amount of data and retain reasonable control of the hits that entered our counts, we picked five more or less evenly spaced decades from which we retrieved the adjectives under investigation.<sup>7</sup> In most cases, attaching the part-of-speech tag for adjectives [j\*] proved useful to enhance the precision of the search. However, for the item *fast*, which was frequently mistagged, all ambiguous tags that incorporated the adjective tag [j\*] among their options were included in the search. For *closed*, which is highly ambiguous between past participles, past tense forms and adjectives, and *shut*, which in addition serves as the present tense and infinitive, it was found that part-of-speech tagging was extremely unreliable. To obtain the full range of results, all instances marked with adjective [j\*] and past participle [v?n\*] tags (in the case of ambiguities, including all tags containing either [j\*] or [v?n\*]) were retrieved.

Further, the hypothesis proposed in the introduction distinguishes not only between mono- and disyllabic adjectives, but also between hits preceding a stressed syllable within a noun phrase and such where this is not the case. In order to isolate rhythmically critical examples of the former type, every corpus search was followed up with another search where the adjective in question preceded a word tagged as a noun [nn\*].<sup>8</sup> On this basis, the percentage of prenominal occurrences (defined above as “critical

---

<sup>6</sup> Viewed from a historical perspective (cf. OED, s.v. *impeccable* adj. and *flawless* adj.), this split might go back to the original uses of these adjectives. *Impeccable* is first attested in 1531 with the meaning ‘of persons: not capable of or liable to sin; exempt from the possibility of sinning or doing wrong’. This meaning is last attested in 1849; from 1620 onwards it developed the meaning ‘of things: faultless, unerring’. *Flawless*, on the other hand, is first attested in 1659 with the meaning ‘free from flaws; without a crack, defect, or imperfection’. It seems to have never referred to persons. While the meanings can be seen to partly converge from 1620 onwards in ‘faultless, without imperfection’, *flawless* (still) cannot collocate with a noun relating to persons, such as *behaviour*.

<sup>7</sup> For details of corpus structure, see the link in the reference section. As indicated by the tabular overview available from that website, our earliest decade (1820-29) contains no newspaper data and our second decade (1860-69) contains a smaller share of newspaper data than the later decades. However, while the normalized frequencies of the adjectives under investigation vary between the genres (for extralinguistic reasons), sporadic checks on our data revealed no influence of genre on the rhythmic effects targeted in this study.

<sup>8</sup> While this search retrieves virtually all instances of the item in question followed by a common noun, it misses other potentially relevant hits that also involve attributive uses preceding a stressed syllable, where this syllable belongs to an adjective, nominalized numeral or proper noun (e.g. *a wealthy Dutch family*, COHA 1999; *the wealthy few*, COHA 1995; *a wealthy Prague family*, COHA 1992). In a sample of 100 hits for the adjective *wealthy*, 46 hits preceded a common noun and thus were correctly retrieved by our search for critical contexts, but there were another potentially relevant 12 hits of the three abovementioned types that escaped the search. The recall for the search for critical contexts was thus  $46/58 = 79\%$ . Conversely, of a random 100 hits retrieved by a search for *wealthy* immediately preceding a noun, 13 preceded a noun that was not initially stressed: The precision of the search for critical contexts was 87% and thus not significantly different from the average of 85% of noun tokens with initial stress. Precision and recall rates vary between adjectives as well as corpus sections. The analyses in section 4 will show to what extent rhythmic effects remain stable in the face of these error margins.

contexts”) out of all occurrences of the adjective in question was calculated as the crucial parameter of comparison in this study.

For both types of contexts, irrelevant hits had to be excluded. For example, due to ambiguous tagging, the search for *fast* retrieved more instances that represented the adverbial use than the adjectival one, and among instances of the latter numerous examples that did not mean ‘quick, rapid’ but ‘firm, constant, steadfast’ (as in *fast friend*, *hard and fast rule*, etc.). Moreover, some of the hits preceding nouns were not adjectival premodifiers of the nouns in question. Thus, manual checks of all hits (in the case of low numbers) or of (one or occasionally two) random samples of 100 examples were carried out to assess the quota of irrelevant hits, which was then, by way of extrapolation, subtracted from the total numbers of hits in that corpus subsection. To exemplify this procedure, the search for the ambiguous item *closed* tagged as an adjective or past participle yielded 2436 hits for the decade 1990-1999. A sample of 100 of these was checked, producing a relatively high proportion of 33 hits (= 33 %) that were irrelevant, i.e. verbal rather than adjectival uses (e.g. *They must have closed their windows ...*, COHA 1997). Further, the search for the same item immediately followed by the tag for a noun produced 429 hits, 100 of which were manually checked, yielding an error rate of 3 % (e.g. *He ... hung the CLOSED sign from the hook*, COHA 1997; *they had quietly closed ranks against him* COHA 1993). Thus, the crucial parameter *percentage of prenominal use* was calculated as  $(429 \cdot 0.97) / (2436 \cdot 0.67) = 416/1632 = 25.5 \%$  (see Figure 5 below).

In an extension of the results to Present-Day English retrieved from COHA, in section 4.5 the same search procedures which were applied to COHA were repeated for the *Corpus of Contemporary American English* (COCA; 450 million words from the period 1990-2012 at the time of retrieval, May 2015). The results will allow us to compare written and spoken usage. While the spoken component (20 %) is considerably smaller than the written one, the hits for the adjectives investigated are plenty in both the spoken and written sections.

These searches, sampling procedures and extrapolations produce highly reliable results where overall numbers of occurrence are not too high and part-of-speech tagging produces low error rates. For those adjectives that are highly frequent and often mistagged, samples of 100 hits per subcorpus and adjective can provide indications as to the direction of the contrast, but larger samples would be necessary to draw firmer conclusions. What can partly make up for high error rates is the large number of independent samples that have been drawn (at least 100 hits per adjective per decade in COHA and per corpus section in COCA). To check the statistical significance of differences observed between pairs of synonymous adjectives, we will thus report results of chi-square tests in footnotes for every single corpus subperiod. The tests calculate the error probability if we reject  $H_0$ , that is if we assume that the rhythmic shape of the adjectives analysed does have an influence on their use in different syntactic functions. Where our samples of prenominal occurrences yielded an error rate of 0 %, chi-square results for the proportion of prenominal uses per adjective will be indicated straightforwardly. Where the samples produced error rates (i.e., occurrences of the search word preceding the POS tag [nn\*] that did not represent actual attributive adjectives), these will be reported as percentages and the chi-square results for the extrapolated proportions of prenominal uses will be indicated in brackets. Since the chi-square results cannot be taken at face value after extrapolation, we will thus adopt a careful interpretation. In total, the cumulative evidence from all case studies we have performed so far points in the same direction so that the findings seem to mutually support each other.

In section 4.3 we extend the diachronic study by a case study on the development of *glad tidings* vs. *happy tidings* and *glad news* vs. *happy news*, based on the corpus *Early*

*English Books Online* (EEBO; version of July 2016). This collection includes “virtually every work printed in England, Ireland, Scotland, Wales and British North America and works in English printed elsewhere from 1473–1700” (<http://eebo.chadwyck.com/about/about.htm>) and covers 130,000 titles and more than 17 million scanned pages. In the EEBO corpus we searched for absolute numbers and relative percentages of competing word-form tokens, including spelling variants.

#### 4. Analysis and discussion

The four case studies to be presented in this section each focus on one particular synonym pair or triplet. Since lexical items – even such belonging to the same grammatical class – differ vastly from each other in their formal, semantic and collocational characteristics, each of the following subsections will have a different focus. Thus, the contrastive study of *rich* and *wealthy* will deal with differences in the denotational range of the adjectives. The comparison of *fast*, *quick* and *rapid* will include some notes on the lexicalization of collocations and on stylistic levels. The analysis of *glad* and *happy* will address the historical dimension of collocations and the question of how to weigh collocational fixations against free choice of synonyms. Finally, the investigation of the two monosyllables *shut* and *closed* will allow us to extend the hypothesis to effects of syllable duration. Far from being exhaustive, these alternating perspectives will serve as mere pointers to the various influences involved in the choice of synonyms.

##### 4.1 *Rich – wealthy*

The usage note on synonyms in *Oxford Learner’s Dictionaries Online* (s.v. *rich*, *wealthy*) defines *rich* in the sense relevant here as follows: “(of a person) having a lot of money, property or valuable possessions; (of a country or city) producing a lot of wealth so that many of its people can live at a high standard”. The definition of *wealthy* is simply phrased in terms of its synonym: “rich”. In addition, the dictionary site answers the question “rich or wealthy?” by stating that “[t]here is no real difference in meaning between these two words.” It furthermore notes that *wealthy* is less frequent than *rich* and is not interchangeable with the latter in “fixed phrases” such as *stinking/filthy rich* or *the rich and famous*. The assumption of synonymy is corroborated by a large number of comparable examples from our corpus searches where the choice of adjective appears more or less random (bold emphasis in all examples is ours):

- (5) a. Are the **rich péople** just going to keep those loaves to grow mold? (COHA nonfiction, 1998)
- b. There are very, very **wéalthy péople** in the neighborhood and some who are just regular old middle class folks ... (COHA news, 1992)
- (6) a. Westerners have heard many times that Shanghai has grown **rich** and adopted Western culture ... (COHA news, 1997)
- b. Market studies suggest that 10 percent of Hungarians have grown **wéalthy** under capitalism ... (COHA news, 1994)

Interchangeability can generally be posited in collocations with humans (e.g. *people*, *kids*, *girl*, *widow*) and collectives (e.g. *countries*, *city*, *family*). However, *rich* (but not *wealthy*) in the corpus data also combines with a wider range of collocates (e.g. *soil*, *literature*, *fabric*, *flavour*, *history*, *resource*, *experience*), which exemplify its wider denotational range.

Predictably, our corpus searches confirm the lower frequency of *wealthy*, which – despite substantial gains in frequency – is still outnumbered by *rich* three- to fourfold in the early 21<sup>st</sup> century. In view of our research hypothesis, however, the more interesting

comparison is between the relative percentages of prenominal uses of both adjectives: If the null hypothesis is to be rejected, examples like (5a), where *rich* premodifies and immediately precedes a noun, should make up a smaller percentage of its total occurrences than examples like (5b), where *wealthy* occurs prenominally and thus resolves the rhythmically critical situation.

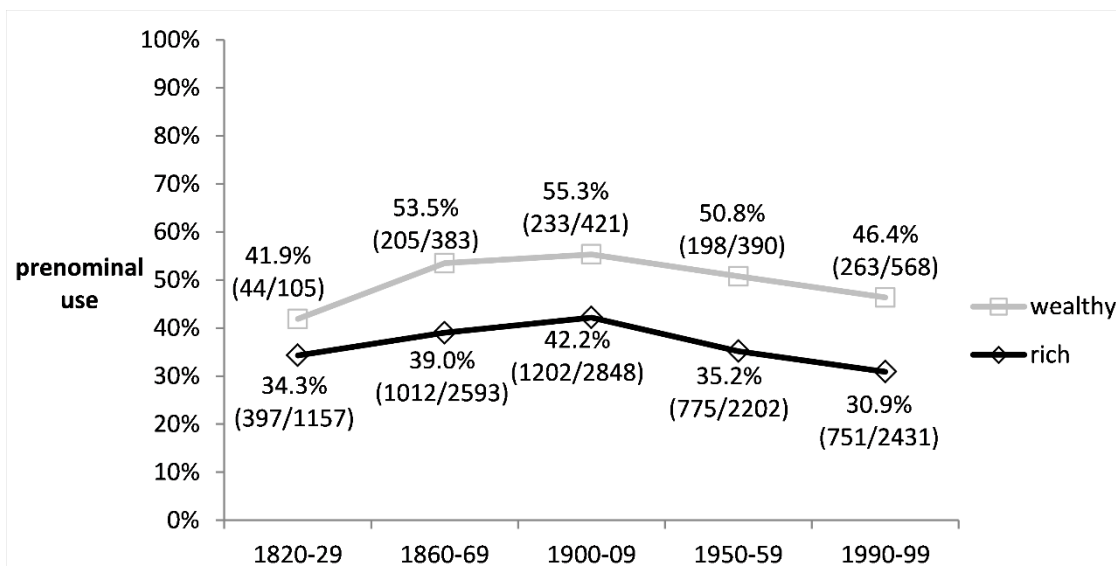


Figure 1: The distribution of *rich* and *wealthy* in COHA, selected decades

Figure 1 presents the results of this analysis. For each decade investigated, the share of instances occurring in rhythmically critical positions (5a/b) out of the total number of instances is indicated. For instance, *wealthy* occurred 105 times in total in the COHA section for the 1820s and all occurrences were examined. 44 of these (41.9 %) were in prenominal, i.e. rhythmically critical positions. In the same subperiod, the adjective *rich* occurred 1157 times in total. Of these, 397 hits (34.3 %) preceded a word that was tagged as a noun, of which 100 were manually checked. As the random sample contained no spurious example, the error rate was determined to be 0 and, by extrapolation, the number of 397 hits was not adjusted. The graph for *rich* demonstrates that this adjective has of course always been used attributively; actually, attributive uses have constantly accounted for 30 to 42 % of the occurrences. Needless to say, attributive use is one of the canonical positions for adjectives, irrespective of their length and despite the fact that with monosyllables this position regularly provokes stress clashes with the following nouns, as in example (5a). However, as Figure 1 shows, the synonym *wealthy* displays a consistently stronger affinity with prenominal uses, occurring in such positions in 42 to 55 % of its occurrences. The difference between the adjectives is statistically highly significant for each of the subperiods investigated, except the first, where *wealthy* is still too infrequent.<sup>9</sup> In line with what was found for adjectives in general (cf. Schlüter, 2005, p. 330; cf. also Cutler & Carter, 1987), both *rich* and *wealthy* are followed by initially stressed nouns in between 84 and 95 % of their attributive occurrences across all subperiods. Considering the substantial synonymy of the adjectives, the distributional difference displayed in Figure 1 can thus be attributed to the difference in their rhythmic

<sup>9</sup> For the last subperiod, we have to factor in a certain sampling error, which is why the  $\chi^2$  result is only given in brackets. The significance levels for the differences in proportions of prenominal use between *rich* and *wealthy* are: 1820-29:  $\chi^2 = 2.44$  (n.s.), 1860-69:  $\chi^2 = 29.01^{***}$ , 1900-09:  $\chi^2 = 25.71^{***}$ , 1950-59:  $\chi^2 = 34.27^{***}$ , (1990-99:  $\chi^2 = 48.86^{***}$ , error rates: *rich* 8 %, *wealthy* 1 %).



shapes: *Wealthy* is more apt to serve in prenominal position because its disyllabic form effectively prevents the adjacency of stressed syllables.

A further hypothesis that will not be pursued here is that in its other senses, *rich* in prenominal position might be partially replaced by alternative adjectives that fit the rhythmical conditions better. For instance, in its sense ‘full of variety’, *rich* might yield to *varied*; when referring to food, *nutritious* or *luxurious* may express the same meaning; when referring to soil, *fertile* is an option; when describing colours or sounds, *vivid*, *vibrant* or *resonant* could be substituted; in yet other contexts, *gorgeous*, *splendid*, *elegant*, *expensive*, *precious* or *priceless* may step into the breach prenominally.<sup>10</sup>

#### 4.2 *Fast – quick – rapid*

As in the case of *rich* and *wealthy*, for *fast*, *quick* and *rapid* corpus examples can easily be found that suggest a potential synonymy of the three adjectives in question.

- (7) a. ... an inferior force can, through **fást móvément**, surprise attack, and intelligent use of the terrain, mount a successful campaign ... (COHA magazines, 1993)  
 b. ... she saw, or thought she saw, the same **quíck móvément**, that hesitation, that almost involuntary act of pulling back from something. (COHA fiction, 1996)  
 c. In addition to an anthropomorphic head-eye system capable of **rápíd móvément**, the Waseda teams are combining voice recognition software and speech synthesizers ... (COHA magazines, 1993)
- (8) a. She never felt the blow, it was too **fást**. (COHA fiction, 1999)  
 b. One’s rapprochement is never as **quíck** and as charming as it is when one is en route. (COHA magazines, 1997)  
 c. When the buildup began, the Pentagon tried to make it appear more **rápíd** than it actually was, ... (COHA magazines, 1990).

However, a quantitative analysis of the triplet incurs a number of difficulties. The adjective *fast* is particularly problematic since its use in the sense of ‘rapid’ developed in late Middle and Early Modern English from the homonymous adverb illustrated in *stand fast*, *run fast* etc. (cf. OED, s.v. *fast*, adj.). Its earlier sense, still well represented in all decades of our corpus though with declining frequency, is that of ‘firm, constant, steadfast’. While the automatic tagging supplied by COHA is often ambiguous between adverb and adjective, this disambiguation as well as that between the two senses of the adjective was easy to perform when inspecting the random samples.<sup>11</sup> In contrast, *quick*, which can likewise function as an adverb, is rarely mistagged in the data, and *rapid* is wholly unproblematic.

To complicate matters further, co-occurrence preferences limit the interchangeability of *fast*, *quick* and *rapid*. *Oxford Learner’s Dictionaries Online* (s.v. *fast*, *quick*, *rapid*) discuss the three in a note labelled “Which Word?” and explain differences of usage primarily in terms of the nouns with which they collocate. In fact, our data demonstrate that towards the end of the 20<sup>th</sup> century, *fast* becomes increasingly involved in a number of fixed collocations (phraseological units, cf. section 4.3). In the 1950s,

<sup>10</sup> The examples are partly taken from Thesaurus.com (2016).

<sup>11</sup> Due to the large number of irrelevant hits that are actually adverbs or adjectives representing the older meaning not synonymous with *quick/rapid*, the high error rates for *fast* lead to considerable insecurity when extrapolation to hundreds or thousands of hits is based on samples of only 100 hits. This insecurity also prohibits an application of the chi-square test to the figures for *fast*.

frequent reference is made to the *fast breeder* (a type of nuclear reactor), while in the 1990s, three collocations alone (*fast food*, *fast lane*, *fast track*) make up around two fifths of the prenominal occurrences (and well over half in spoken American English, which will be sketched in section 4.5). The first of these recent phraseological units may represent a deliberate formation involving an attention-grabbing alliteration plus stress clash (*fást fóod*).<sup>12</sup> The latter two even show evidence of being compounds, with their main stress on the initial element (*fást lànè*, *fást tràck*).<sup>13</sup>

Thus, we have to allow for various distorting factors impinging on the distribution of the three adjectives studied. However, there is no *a priori* reason to expect that they should differ in the syntactic positions they take – if not the rhythmic shape that sets the disyllable *rapid* off against the monosyllables *fast* and *quick*. Figure 2 provides further evidence for this research hypothesis.

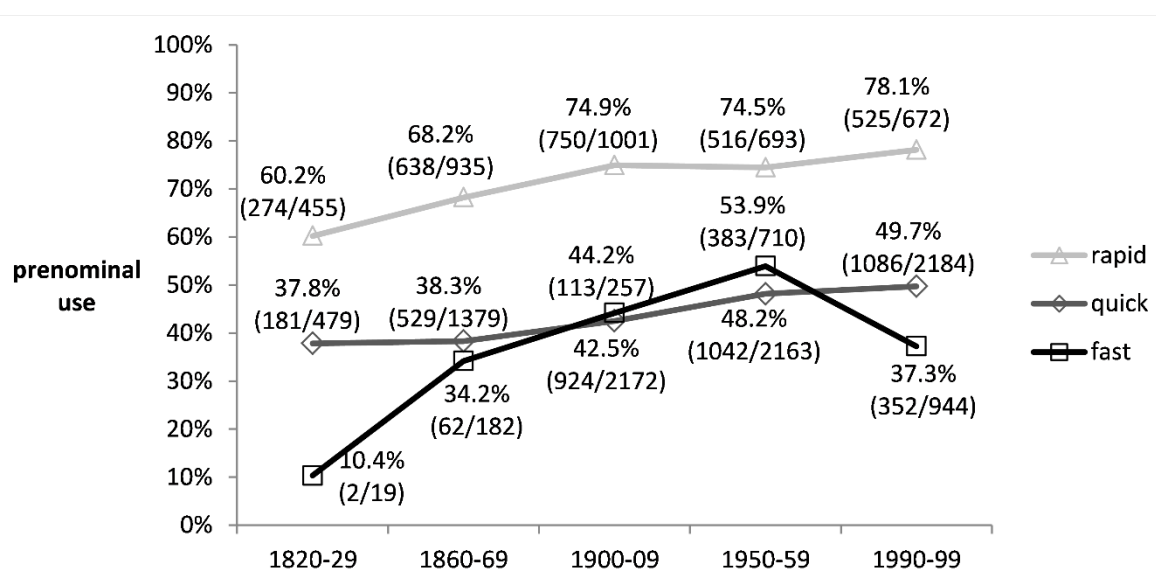


Figure 2: The distribution of *fast*, *quick* and *rapid* in COHA, selected decades

In the earliest subperiod, *fast* is generally very rare as an adjective in the sense of ‘rapid’, and in the later subperiods, it produces massive error rates. Yet, visual inspection of the results for *fast* and *quick* on the one hand and *rapid* on the other, and statistical testing of the distributional contrast between *quick* and *rapid* reveal a diachronically stable and statistically convincing contrast: *Fast* and *quick* appear in rhythmically critical positions between 34 and 54 % of their total occurrences, while *rapid* figures extremely often prenominally.<sup>14</sup> This finding supports the hypothesis that the syntactic distribution of synonymous adjectives is at least partly governed by the avoidance of stress clashes.

An inspection of the random samples for *rapid* however leads to a supplementary explanation: This adjective, borrowed from Latin or French as late as in the 17<sup>th</sup> century

<sup>12</sup> See section 4.3 for a consideration of the possible poetic function of such phraseological units.

<sup>13</sup> As before, primary lexical stress is indicated by acute accents on the syllable nuclei. The more reduced secondary stress usually found in the non-initial component of compounds is symbolized by grave accents (Compound Stress Rule; cf. Hayes, 1995, p. 368). The stress assignment in the given phraseological units and compounds follows the EPD (s.v. *fast*). This raises the question of whether *fast lane* and *fast track* should be excluded from the counts on account of being compounds, which would reduce the number of prenominal uses of *fast* substantially for the 1990s and reinforce our conclusions.

<sup>14</sup> Comparing the shares of prenominal uses between *quick* and *rapid*, we find low error rates and thus significant distributional differences throughout: (1820-29:  $\chi^2 = 47.00^{***}$ , error rates: *quick* 2 %, *rapid* 0 %), (1860-69:  $\chi^2 = 198.93^{***}$ , error rates: *quick* 1 %, *rapid* 0 %), (1900-09:  $\chi^2 = 288.31^{***}$ , 1950-59:  $\chi^2 = 146.26^{***}$ , 1990-99:  $\chi^2 = 168.560^{***}$ ).

(cf. OED s.v. *rapid*, adj., adv. and n.) is not only the preferred choice to characterize the speed of processes of change (rather than of movements or actions; cf. *Oxford Learner's Dictionaries Online*, s.v. *rapid*), but it often premodifies nouns of Latinate origin, too (e.g. *ráte, páce, chánge, devélopment, evolútion, succéssion, expánsion, impróvement, réaction, demíse, transítion*). These nouns often have abstract meanings, belong to rather formal registers and, crucially, many of the nouns found in the samples lack the initial stress typical of Germanic words. Compared to the average of nouns in English, 85 % of which are initially stressed, noninitially stressed nouns are thus massively overrepresented after *rapid* (21-37 % in COHA; 38-45 % in COCA). While this might lead us to expect a greater compatibility with *fast* or *quick* (which is not borne out by the data), stylistic and genre-specific affinities obviously have to be allowed for in accounting for the co-occurrence preferences of *rapid*.

In sum, while this case study has unearthed further factors impacting on the choice of adjectives in prenominal positions (the establishment of preferred collocations, their fixedness as compounds and stylistic attraction), the extraordinarily high incidence of *rapid* before nouns can still be ascribed to the Principle of Rhythmic Alternation: The majority of nouns following *rapid* in our data are, after all, initially stressed, so that the use of *rapid* helps to avoid stress clashes. For the minority of nouns that are not initially stressed, we might consider a carry-over effect from the larger number of rhythmically critical cases: *Rapid* is the one adjective among those investigated in this paper that is most strongly attracted to prenominal contexts, for which it shows a noticeable predisposition.

### 4.3 *Glad - happy*

In contrast to the adjectives discussed so far (and *rapid* in particular), the pair *glad* and *happy* (as well as *shut* and *closed*, investigated in the next section) are substantially less frequent in attributive uses. Gladness and happiness in human beings (as well as the condition of being shut or closed in objects that can also be open; see below) tend to be transitory states. Thus, a closer look at the semantics of the adjectives under investigation and concomitant syntactic effects is in place.

Adopting a semantic approach, Leisi (1985, p. 54) argues that the meaning of *glad* is restricted to actual uses, i.e. to cases where the referent is currently in a state of gladness. This semantic feature is mirrored in the syntax of *glad* (with the exception of a few fixed collocations, on which see below): It can, according to Leisi, only be used in predicative positions, but not in attributive positions, which typically characterize the referent of the noun permanently. In contrast, *happy* is subject to no such restrictions; like the majority of other adjectives, it has actual and potential uses alike, meaning that it can also refer to a person in whom happiness is a trait of character, even though at the moment of speaking he or she may be in a less-than-happy mood. As a consequence, *happy* should occur unrestrictedly in predicative or attributive position (on semantic nuances of attributive and predicative adjectives, cf. also Bolinger, 1952, p. 1133-1137; 1967, p. 3-4; Jacobsson, 1996, p. 218). An inherent semantic difference between *glad* and *happy* is thus predicted to lead to the same distributional tendency as rhythmic considerations. Both motivations conspire to disfavour *glad* in prenominal position, and on both accounts *happy* can be expected to occur in prenominal as well as other (predicative and postnominal) positions.

A full semantic analysis of data samples is beyond the scope of the present study. However, the following examples have been selected so as to show that *glad* and *happy* do not necessarily occur in contexts forcing a different semantic interpretation, and examples like these are numerous.

- (9) a. Sir John Maynard moved to make bonfires all over the town, and the House rose so that members could spread the **glád néws**. (COHA nonfiction, 1996)  
 b. We were housed in the Palace as the chief envoy went to carry the **háppy néws** to the Prince. (COHA fiction, 1997)
- (10) a. Her mother was still talking with Marla's, who seemed **glád** to see Janice. (COHA fiction, 1999)  
 b. People seemed **háppy** to be out with their children, gazing at animals, drinking soda, ... (COHA fiction, 1998)

Especially in the earlier decades, both adjectives frequently premodify nouns denoting time or temporally limited events, such as *time, hour, day, morning, moment, laughter, smile, songs, cries, voice, tidings, news, memories* or *heart* ('feeling'), in which Leisi's (1985, p. 54) distinction between actuality and potentiality is inapplicable: The temporal extension of gladness/happiness coincides with the temporal extension of the referent: A particular piece of *news* is characterized as *glad*, an instance of *laughter*, while it sounds, is *happy* throughout, and so on. In contexts like these, *glad* and *happy* seem to be interchangeable, at least to a certain degree. However, it is true that *happy* also occurs as an attribute before nouns designating persons, where *glad* appears extremely rarely, for instance before *child, couple, face, family, man, people* and *person*.

Espousing Bolinger's (1965b, p. 149) presupposition that *glad* and *happy* are interchangeable in at least some of their uses, the analysis presented in Figure 3 follows up his hunch that *happy* may replace *glad* in rhythmically critical positions. As in the previous case studies, we will test whether *happy* occurs more frequently prenominal than *glad*.

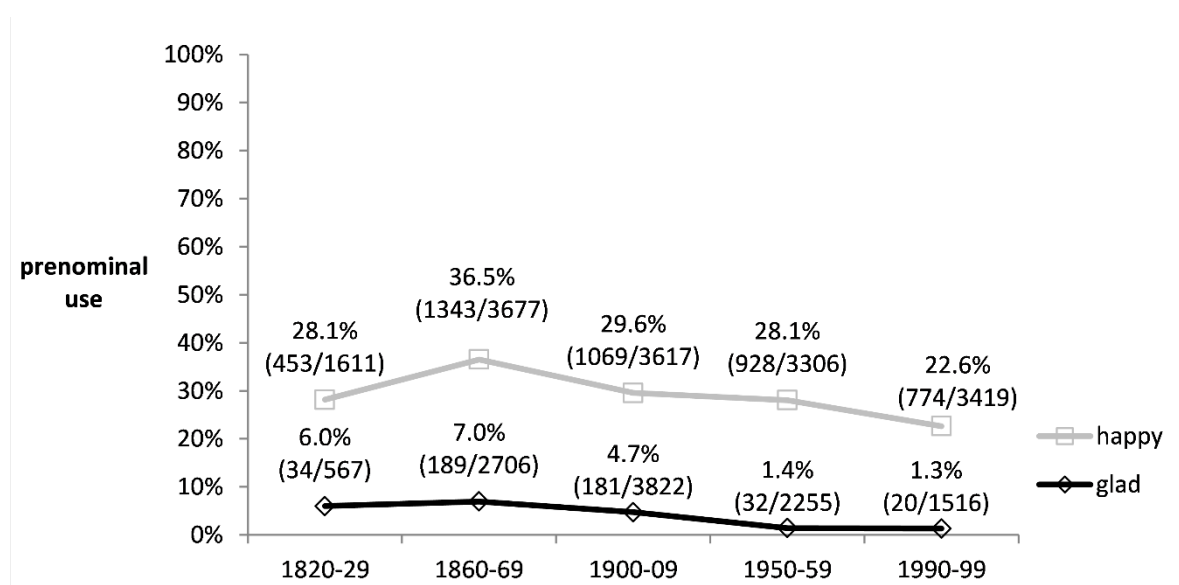


Figure 3: The distribution of *glad* and *happy* in COHA, selected decades

The research hypothesis seems to be borne out by the data: For every decade, *happy* occurs clearly more often in prenominal position than *glad*.<sup>15</sup> At the same time, a further

<sup>15</sup> The chi-square results are only slightly kerbed by small error rates: (1820-29:  $\chi^2 = 118.23^{***}$ , error rates: *glad* 0 %, *happy* 1 %), 1860-69:  $\chi^2 = 745.72^{***}$ , (1900-09:  $\chi^2 = 818.84^{***}$ , error rates: *glad* 0 %, *happy* 9 %), 1950-59:  $\chi^2 = 666.67^{***}$ , (1990-99:  $\chi^2 = 353.58^{***}$ , error rates: *glad* 0 %, *happy* 2 %).

decline in the already low numbers of prenominal use of *glad* is noticeable. In the 19<sup>th</sup> century, *glad* seems to have been less problematic in prenominal use than later. This leads us to a subsidiary hypothesis, the testing of which is beyond the scope of this study since it would require a qualitative (semantic) analysis of a large number of examples reaching back further than the earliest COHA decade: We propose that the rhythmic inappropriateness of *glad* in prenominal position and its increasing restriction to predicative and postnominal uses may in its turn have led to a gradual loss of potential (characterizing) meanings in the adjective. In other words, rhythmic constraints may have given rise to an ensuing semantic restriction of *glad*, which never affected *happy* thanks to its different rhythmic shape and continued prenominal occurrence. While the emergence of such a secondary semantic specialization has – to our knowledge – never been supported with empirical data, there is the comparable case of so-called “predicative-only” adjectives (*alive, asleep, adrift, afraid, aware* etc.), which have been argued to be generally barred from attributive uses on account of both their semantics and rhythm (cf. Bolinger, 1965b; Jacobsson, 1996; Schlüter, 2008).

The most frequent collocation in our COHA and COCA returns for prenominal uses of *glad* is *glad tidings*. A closer look at this collocation leads us to a consideration of the role of (‘fixed’) collocations in our study. In particular, we will study the combination *glad* + noun from two points of view. First, we will show in a deeper diachronic perspective how *glad tidings* developed from the rhythmically uncritical collocation *glad(d)e tiding(e)s* to the modern collocation with clashing stresses. Second, a look at more modern phraseological formations in our results, which appeared with a stress clash right from the start, such as *glad eye*, will lead us to a consideration of the potential impact of the poetic use of language on our findings. In the following, the cover term ‘phraseological unit’ will specifically refer to adjective + noun collocations that are habitually employed to express a given idea, and as such form a semantic unit.

*Glad tidings* is a well-known example of a phraseological unit. In Cowie, Mackin and McCaig’s (1983) idiom dictionary, *(the) glad tidings* is labelled ‘formal’ and defined as ‘very good news’ (cf. also OED, s.v. *glad* adj. 4). Among our few corpus returns for *glad* in prenominal use, *glad tidings* occurred most frequently. High string frequency may seem to be an obvious criterion of phraseological status, but we would argue that it is not a necessary criterion. Thus, the phraseological unit *glad eye* ‘a look or movement of the eyes designed to attract a person of the opposite sex’ (cf. OED s.v. *glad* adj. 4) is much less frequent (only 2 occurrences in COCA) than *glad tidings* (33 occurrences), and also less frequent than the free combination *glad cry* (9 occurrences). Identifying phraseological units with the help of corpus linguistic means is a problematic issue: Taken on their own, neither String Frequency, nor the measures of Transitional Probability, T-scores or (Pointwise) Mutual Information help to reliably find these units.<sup>16</sup> Due to the lack of a workable exclusion measure for phraseological units with prenominal *glad*, we decided to adopt an inclusion strategy for practical purposes and for the general survey in Figure 3. As has been previewed above, it is worthwhile to study two probable phraseological units separately to exemplify possible reflexes of phraseological fixations. Our assumption is that the phraseological units *glad tidings* and *glad eye* constitute two different types of violation of the Principle of Rhythmic Alternation, which both result from their phraseological status. One is the retention of a phraseological unit despite a newly developing violation of rhythmic alternation (*glad tidings*); the other concerns

---

<sup>16</sup> Cf., e.g., Roth (2014, p. 28-29, 66, 71), who employs a combination of these measures together with Saliency for his analysis. Cf. also the overview in Xiao (2015, p. 108-112), and cf. Gray and Biber (2015) on possibilities of corpus-driven analyses of phraseology for Present-Day English.

phraseological units which have embraced the rhythmical prominence of two clashing monosyllables since their creation (*glad eye, glad hand, glad rags*).

The EEBO database lists different spellings of *glad tidings* in the 16<sup>th</sup> and 17<sup>th</sup> centuries. In Middle English, all adjectives were suffixed with *-e* (schwa) except those which were monosyllabic, ended in a consonant, had singular reference and were not preceded by a demonstrative pronoun, a possessive or the definite article (i.e. formerly strong adjectives, mostly in predicative position). Minkova (1991, ch. 7) suggests that the prolonged retention of *-e* in monosyllabic singular weak adjectives – and likewise in the plural (cf. 1991, p. 186) – goes back to the rhythmic preference of stress-clash avoidance in adjective + noun combinations. Thus, in the earliest attestation of *glad tidings* in the OED, *-e* indicates a rhythmically harmonious phrase:

- (11) Let him in..he bringeð us **gleade tidinges**. (*Sawles Warde* in *Cott. Hom.* 257, a1240; OED, s.v. *glad* adj. 4b)

But by c. 1450 the last schwas in final position had disappeared, and there are no indications that final *-e* was still pronounced in the educated speech of London after the early 15<sup>th</sup> century (cf. Dobson, 1968, p. 879).

A fifteenth century reversal of the strength of the eurhythmy rules relative to other rules would not [sic] surprising. Extremely powerful morphological analogy within the adjectives, demorphologization of *-e* in the other word classes, as well as sweeping changes in the syntactic structure of the language led to what may appear to be a tendency to “dysrhythmy” in the prosodic organization of monosyllabic adjective + initially stressed noun phrases in Modern English. (Minkova, 1991, p. 180)

One strategy to avoid this new stress clash was to substitute the disyllabic synonym *happy* for the now monosyllabic *glad*. This strategy is mirrored by the situation in Present-Day English described above (cf. Figure 3). However, as Figure 4 shows, with c. 8 % of all occurrences in the 16<sup>th</sup> and c. 6 % in the 17<sup>th</sup> century, *happy tidings* remained the rarer option, and the phraseological unit *glad tidings* was retained. The noun *news*, which developed as a synonym of *tidings* from the 15<sup>th</sup> century onwards (cf. OED, s.v. *news* n. 2a), also entered into a collocation with *glad* (here, final *-e* with only 2 out of 158 instances in EEBO was not extended to the new expression), but it can be seen that the rhythmically more satisfying collocation *happy news* was preferred over *glad news* in the first centuries of use (80 % *happy* in the 16<sup>th</sup> century, 66 % in the 17<sup>th</sup> century) and is again preferred in Present-Day English (c. 94 % in COCA). As Figure 4 shows, however, *glad news* acquired some popularity in the 19<sup>th</sup> century (maybe inspired by the model of *glad tidings*), but the COCA data reveal that there was a steep decline again at the end of the 20<sup>th</sup> century in favour of the combination *happy news*, without stress clash.<sup>17</sup> In Present-Day English, thus, the rhythmically marked phraseological unit *glad tidings* ultimately prevailed over its non-phraseological competitor *happy tidings*, while at the same time its non-phraseological synonym *happy news* with regular rhythm ousted the rhythmically marked but not phraseologically fixed collocation *glad news*.<sup>18</sup>

<sup>17</sup> 1500-1600:  $\chi^2 = 201.81^{***}$ , 1601-1700:  $\chi^2 = 943.67^{***}$ , 1810-1899:  $\chi^2 = 5.04^*$ , 1900-89:  $\chi^2 = 27.99^{***}$ , 1990-2015:  $\chi^2 = 76.96^{***}$ .

<sup>18</sup> For an overview of English historical phraseology and avenues for research, see Knappe (2012).

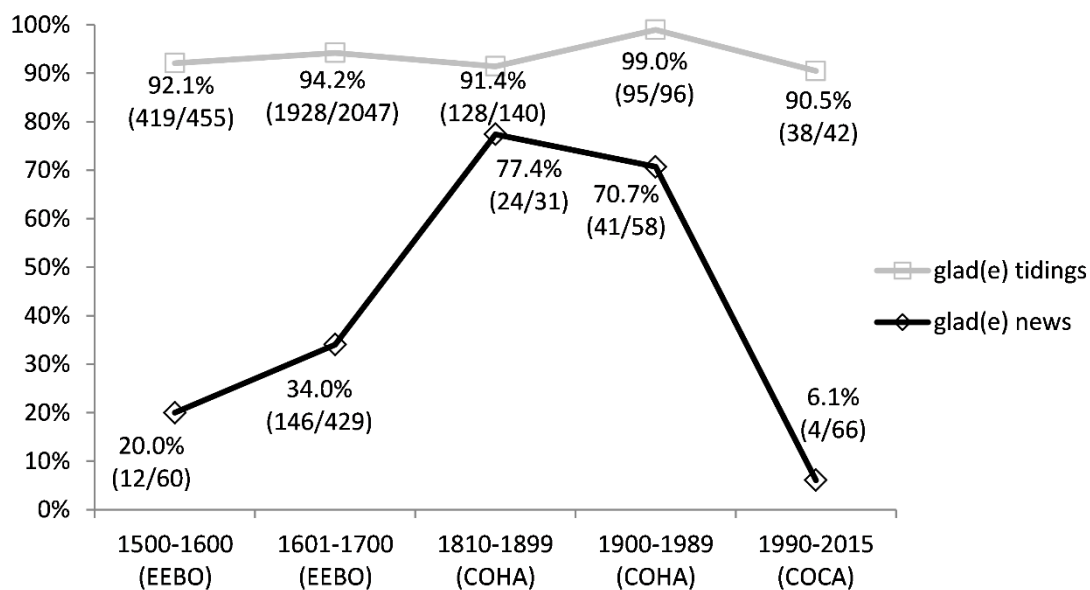


Figure 4: The use of *glad(e)* as opposed to *happy* premodifying *tidings* and *news* in EEBO, COHA and COCA

Interestingly, a manual check of all 1,951 occurrences of *glad(d)e* in EEBO revealed that in comparison with free collocations the adjective was comparatively slow to lose its *-e* in phraseological units: c. 12 % of all instances of *glad(d)(e) tidings* (+ variants) in the 16<sup>th</sup> century (49 of a total of 419 tokens) and still c. 1 % in the 17<sup>th</sup> century (16 of a total of 1,928 tokens) retained *-e*. The following quotations show the difference with regard to *-e* between the phraseological unit *glade tyding(e)s* and the free collocation *good things*, and how the phraseological unit, too, gradually lost orthographic *-e* in the course of the 17<sup>th</sup> century:

- (12) And indeed if we had the sight of the fearcenesse of that eternall wrath of God for sinne, it could not be possible but we would say, *O hovv bevvtifull are the feete of them vvhich bring **glade tydings** of peace, and bring **glade tydings** of **good things**!* (EEBO: Robert Rollock, *Lectures vpon the first and second Epistles of Paul to the Thessalonians ...*, 1606)
- (13) *That faith is by hearing, for which effect preaching is needfull, and therefore he with the Prophet sayeth there, how beautifull upon the mountaines, are the feet of them that bring **glad tydings**, that bring **glade tydings** of **good things**?* (EEBO: James Sibbald, *Diverse select sermons vpon severall texts of holy scripture ...*, 1658)

Most of the other instances of *glad(d)e* + noun also belong to recurring expressions (i.e. potential phraseological units); these are in particular: 1) 'glad facial expression': (*with [a] / of / maken*) *gladde chere / cheres, (with a) gladde vysage, (with a) gladde countenance*, 2) 'glad heart and mind': (*with [a]*) *gladde harte / hartis and mynde, with glade curage*.

Thus it seems that before the final fixation of spelling, phraseological units tended towards a more conservative spelling. Whether this is a sign of a somewhat prolonged time of stress-clash avoidance in these expressions cannot be judged from the sources at hand. In any case, the spelling is indicative of a special status of the unit.

Secondly, while time-honoured fixedness made the collocation *glad tidings* survive its acquired stress clash, newly coined phraseological units can actually be created with a

stress clash right from the start, and then – we will argue – rhythmic marking complements pragmatic marking. The focus here is on three phraseological units which were coined around the turn of the 20<sup>th</sup> century (all OED, s.v. *glad* adj. 4): *glad hand* ‘the hand of welcome, frequently used somewhat ironically’ (first attestation: 1895), *glad rags* ‘one’s best clothes; very smart or fancy clothes; *spec.* formal evening dress. *colloq.* (orig. U.S.)’ (first attestation: 1902), and *the glad eye* (first attestation: 1911).

- (14) The lover, the elderly philanderer, the girl with the glad eye. (*Punch* 22 Nov. 382/2, 1911; OED, s.v. *glad* adj. 4)

Roman Jakobson (1960, p. 356f.), in his famous discussion of the poetic function of language, draws attention to the political slogan *I like Ike*, which achieves its function through the combination of sound structure and sentence semantics (e.g. *Ike* being fully enveloped by *like*). On top of this, one may add, the double stress clash makes the slogan phonologically more salient. While a combination of adjective + noun is certainly a different kind of phraseological unit from a slogan, and although English phraseological units are often metrically well-formed (cf. Naciscione, 2010, p. 37f.), in the case of *glad eye* and *glad hand* (as well as that of *fast food*, mentioned in 4.2) it can be hypothesized that in the process of coining the units, stress clashes were embraced as pragmatic markers.<sup>19</sup> In all three cases, according to the OED (*glad* adj. 4), *glad* denotes ‘filled with, marked by, or expressive of joy or delight’, and thus *joyful* or *delightful* could have been chosen instead of *glad* to produce rhythmically uncritical combinations. However, stress clashes persist, though with different leanings:<sup>20</sup> Phrasal stress seems to be retained in *glád éye* and *glád hánd* (as well as *fást fóod*, mentioned in 4.2), while *glád ràgs* (as well as *fást làne* and *fást tràck*, also in 4.2) has compound stress.<sup>21</sup> Though these examples, viewed superficially, appear to contradict the hypothesis proposed in the present contribution, the tendency to provoke stress clashes in phraseological units like these does not call the efficacy of the Principle of Rhythmic Alternation into question. Rather, the fact that rhythmic alternation can be interrupted for pragmatic marking underlines the importance of the phenomenon. An intentional infraction of rhythmic alternation is an appropriate means of giving the utterance more prominence, exactly because it is an effect that is normally avoided.

#### 4.4 *Shut – closed*

*Oxford Learner’s Dictionaries Online* (s.v. *closed*, *shut*) flatly note that “[c]losed is used in front of a noun, but **shut** is not”. Similarly, Jakobsson (1996, p. 209) remarks that *closed eyes* is fully acceptable as an attributive construction, while *\*shut eyes* is objectionable, although *shut* is largely synonymous with *closed*. Being concerned with explanations that are syntactic or semantic in nature, Jakobsson is unable to account for this phenomenon. An explanation in terms of the preference for alternating strong and weak syllables is not available either, since both *closed* and *shut* are monosyllabic. A clue to this phenomenon

<sup>19</sup> Bolinger (1981, p. 29; 1986, p. 70) surmises that the purpose of such clashes is a vigorous and emphatic reinforcement of what is said.

<sup>20</sup> The stress assignment in the given phraseological units and compounds follows the EPD (s.v. *glad*). Where the EPD lists no special stress pattern, ordinary phrasal stress is assumed, i.e. the adjectival premodifier and the nominal head retain their full lexical stresses (with the stress on the head exceeding that on the premodifier, in line with the Nuclear Stress Rule; cf. Hayes 1995, p. 368).

<sup>21</sup> It is not uncontroversial in the literature on word formation whether adjective + noun combinations such as these started out as units with phrasal stress prior to their compound status (cf. Štekauer, 2000, p. 213, note 6).



is provided by Bolinger (1965a, p. 135).<sup>22</sup> The two adjectives differ greatly in their phonetic duration: While *closed* contains a diphthong as its syllable nucleus and ends in a voiced consonant cluster, *shut* only has a short vowel, whose shortness is moreover reinforced by the voiceless stop in its coda. In consequence, the considerable duration of the monosyllable *closed* offers sufficient phonetic substance to be drawn out so as to accommodate a drop in articulatory energy from a stressed beginning to a much less stressed termination (cf. Bolinger, 1965a, p. 135; Hayes, 1984, p. 72). This effect is similar in nature to that of a second unstressed syllable: If followed by another stressed syllable, a long monosyllable incorporates the part of a buffer spacing out two stressed syllables.<sup>23</sup>

This section will thus put forward a supplementary hypothesis along the following lines: While ideally, two stressed syllables are separated by an intervening unstressed syllable, in the absence of an unstressed syllable, a long syllable in the first position will be preferable to a short one. The predicted outcome of this preference is that in our corpus, *shut* should be noticeably avoided in attributive uses, while *closed* should be more free to occur in such positions.

As pointed out in section 3, the search procedure was far from unproblematic since the automatic tagging often failed to disambiguate instances of *shut* and *closed*. Instances not preceding a noun were frequently tagged as participles (prenominal position obviously being used as a clue to adjectivehood), so that our search for critical contexts as well as for general frequencies of *shut* and *closed* targeted all instances tagged as adjectives or past participles and all random samples were manually sorted into adjectival and adjective-like uses on the one hand and verbal uses on the other. In the process, other irrelevant hits such as the compound *shut eye* 'sleep' and the phraseological unit *open and shut case* 'legal case or other matter that is easy to decide or solve',<sup>24</sup> where *shut* is not a prototypical attribute, were eliminated. Figure 5 provides the results in the familiar format.

---

<sup>22</sup> Bolinger (1965a, p. 135) reports an elicitation experiment employing the pair of nonsensical sentences in (i) and (ii) to elicit preference judgements from consultants. As expected, his consultants tended to prefer example (ii) with the phonetically long monosyllable *plam* in the clashing position to (i) with the considerably shorter *plap* in the same position.

(i) He lives in a **plám** and **pláp hóuse**.

(ii) He lives in a **pláp** and **plám hóuse**.

<sup>23</sup> Compare also the phenomenon known as *raddoppiamento sintattico* frequently described in Italian phonology: A consonant in the onset of the second of two clashing syllables can be lengthened so as to put more distance between the two stressed syllable nuclei (cf. Bayer, 1989, p. 16).

<sup>24</sup> Definitions taken from *Oxford Learner's Dictionaries Online*.

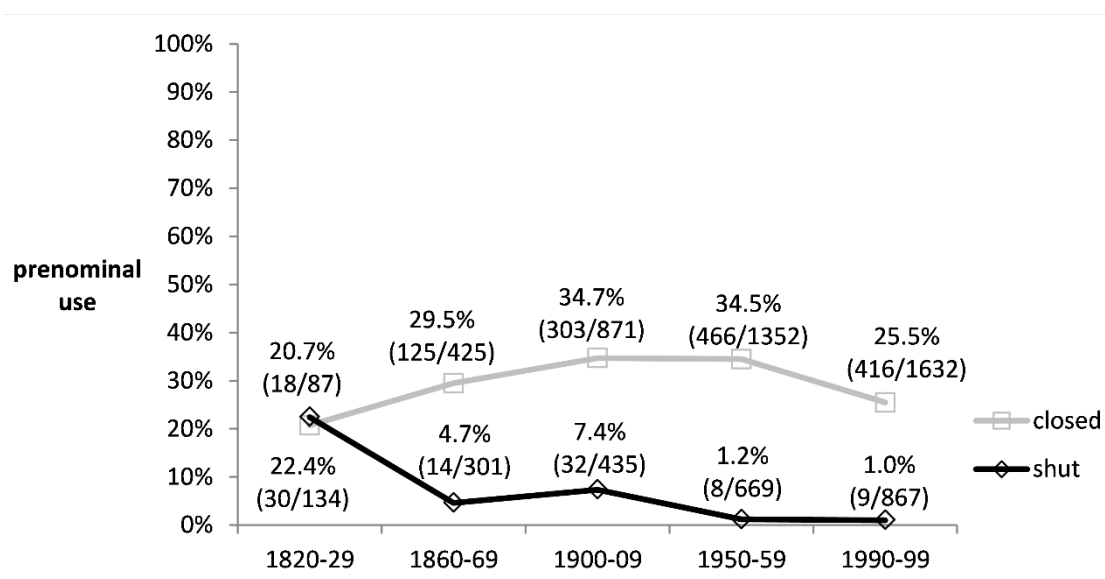


Figure 5: The distribution of *shut* and *closed* in COHA, selected decades

The results of the corpus analysis clearly show a consistent difference in the distribution of the two adjectives, except for the data from the early 19<sup>th</sup> century. *Closed* is used preminally in between 25 and 35 % of its occurrences and thereby contrasts sharply with *shut*, which hardly occurs in this position after the 1820s.<sup>25</sup> Detailed inspection of the sampled hits reveals that in all contexts where *shut* is used, *closed* could be used as well, since reference is typically made to *mouths*, *lips*, *eyes*, *lids*, *doors* and *windows*. The examples in (15) are representative in this respect.

- (15) a. As I floated, the sun a red mist beyond my **shút** **éyelids**, I felt the slow, languid pace of life here ... (COHA fiction, 1995)  
 b. I covered my eyes while my wife read the article sitting on the **clósed líd** of our toilet. (COHA fiction, 1997)  
 c. But at eight o'clock, Daphne came to the **shút** **dóor**, waited a minute, creaked it open and said ... (COHA fiction, 1991)  
 d. Alice glanced at the **clósed** **dóor**, as if fearing Verna could see her. (COHA fiction, 1997)

The same collocates are found with the relatively numerous examples of prenominal *shut* from the 1820s. This raises the question of why the imbalance found in later decades is not yet in place in the earliest subperiod. Part of the answer may lie in the fact that the frequency of *closed* has more than tripled and thereby overtaken that of *shut* in the course of the past two centuries, as is revealed by a quick check in COHA: *Closed* was not as obvious a bypass then as it is now.

Beyond concrete uses such as those in (15), *closed* – but not *shut* – is also found in more abstract senses (describing *ranks*, *meetings*, *hearings*, *societies*, *parties*, *committees* etc.). The denotational range of *closed* is thus larger, but it properly includes that of *shut*, so that *closed* constitutes a viable alternative to *shut* in rhythmically critical contexts. The number of cases in which this (or some other) solution is not resorted to in the data – in

<sup>25</sup> Due to high error rates in the samples for *shut*, the insecurity introduced by extrapolation invalidates the application of chi-square tests. The results are given in brackets here for the sake of completeness: 1820-29:  $\chi^2 = 0.09$  n.s., (1860-69:  $\chi^2 = 69.78^{***}$ , error rates: *shut* 42 %, *closed* 5 %), (1900-09:  $\chi^2 = 114.47^{***}$ , error rates: *shut* 29 %, *closed* 3 %), (1950-59:  $\chi^2 = 275.96^{***}$ , error rates: *shut* 58 %, *closed* 2 %), (1990-99:  $\chi^2 = 239.85^{***}$ , error rates: *shut* 71 %, *closed* 3 %).

other words, the residue of prenominal uses of *shut* – approaches the zero line in the course of the 20<sup>th</sup> century. The observations by Jakobsson (1996, p. 209) and *Oxford Learner's Dictionaries Online* quoted at the outset are thus confirmed and lend plausibility to the rhythmic account suggested by Bolinger (1965a, p. 135): A speaker or writer heading for a construction like (15a) or (15c) might run into rhythmic difficulties and in many cases substitute the short monosyllable *shut* with the longer form *closed* to avoid the imminent stress clash.

Incidentally, it should be noted that the corpus analysis yields some – statistically insignificant – support for the assumption that premodification of *shut* by an adverb might redeem the rhythmic situation to some extent. Thus, in examples like (16), *half* and *tightly* not only modify the meaning of *shut*, but also deflect the main stress from it, leaving it with no more than secondary stress (backshifted stress; cf. Bolinger, 1965b, p. 141-145; Schlüter, 2005, p. 29-30). This, in turn, separates the major stresses in line with the Principle of Rhythmic Alternation, with *shut* itself functioning as a buffer syllable.

- (16) a. He did not see the gleam in Asbury's **hálf shùt éyes**. (COHA fiction, 1904)  
 b. On Oct. 21, the first trainload of oil from China's Kazakh holdings headed toward Chinese refineries across what until recently had been one of the most **tíghtly shùt bórders** in the world. (COHA news, 1997)

By way of a preliminary conclusion, the results from the four diachronic studies presented so far can be summarized as follows: Even though all the lexemes can in principle occur in single or complex attributive positions as well as in non-attributive positions, the rhythmically less appropriate members of the pairs (or triplet) are strikingly underrepresented before nouns. What is more, these differences have been in place throughout the two centuries investigated. The only statistically relevant exception is the earliest decade in the fourth case study, which may simply be explained by the fact that *closed* had not yet developed into an equally viable alternative to *shut*.

#### 4.5 Spoken and written Present-Day English

So far, we have exclusively used the written data contained in COHA, but they have provided consistent evidence in favour of the avoidance of stress clashes. When it comes to spoken data, we might predict a rhythmic effect that should be at least as strong, since infractions of the Principle of Rhythmic Alternation concern the spoken form of language rather than its written form. Thus, the previous four case studies will be repeated based on COCA, which allows us to distinguish between written and spoken data. For these analyses, all the procedures, caveats and restrictions adduced in sections 4.1 to 4.4 apply identically, but will not be repeated here. The focus will be laid on the contrasting distributions of mono- vs. disyllabic and short vs. long monosyllabic adjectives in prenominal position, respectively.

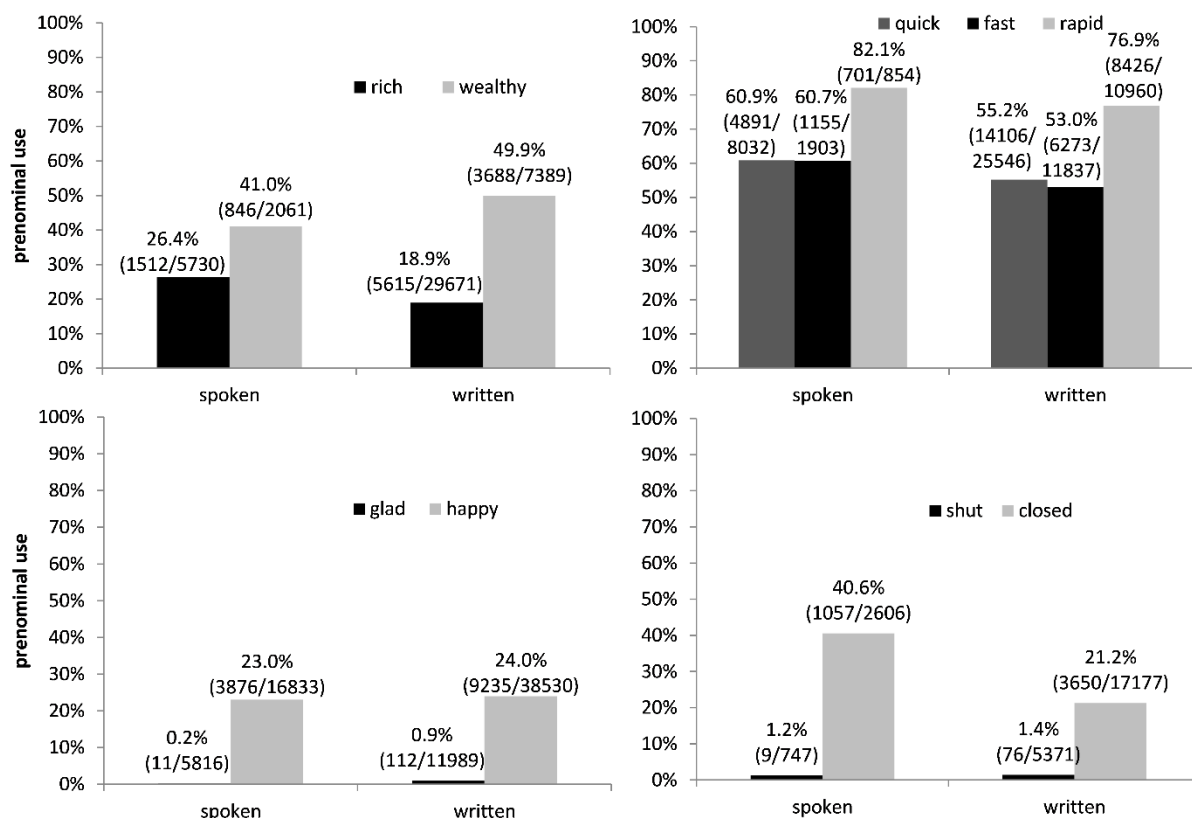


Figure 6: The distribution of (near-)synonymous adjectives in COCA, spoken vs. written sections

The most striking result of the synopsis presented in Figure 6 is that the proportion of cases in which potentially synonymous adjectives are placed in prenominal position differs widely between the triplet *quick/fast/rapid*, all three of which are extremely frequent as noun premodifiers, and the pairs *glad/happy* and *shut/closed*, which have rather low prenominal shares across the board, with *rich/wealthy* taking an intermediate position. As mentioned above, this may be due to their different predispositions to encode actual or potential meanings. Not surprisingly, the distributions in COCA are also very similar to those found in the later subsections of COHA.

Crucially, however, all four panels show that the same contrasts as have been attributed to rhythmic constraints in the previous discussion are observable in the spoken as well as in the written mode: In each case, the disyllabic adjectives and *closed* are used comparatively often in prenominal position, while the proportions of prenominal *rich*, *quick*, *fast*, *glad* and *shut* are substantially smaller. The percentages by which longer adjectives differ from shorter ones are relatively constant across the pairs and the triplet considered: They oscillate between 15 and 30 percentage points, even for the pair of monosyllables that are merely distinguished by phonetic duration.<sup>26</sup> If anything, the contrast is strongest in the spoken data for *shut* and *closed*, which diverge by almost 40%.<sup>27</sup> Without going into further detail here, it may be added that we have obtained

<sup>26</sup> COCA spoken: (*rich/wealthy*:  $\chi^2 = 154.36^{***}$ , error rates: *rich* 8 %, *wealthy* 0 %), *quick/rapid*:  $\chi^2 = 148.58^{***}$ , (*glad/happy*:  $\chi^2 = 1585.70^{***}$ , error rates: *glad* 0 %, *happy* 3 %), (*shut/closed*:  $\chi^2 = 414.68^{***}$ , error rates: *shut* 87 %, *closed* 0 %).

COCA written: (*rich/wealthy*:  $\chi^2 = 3021.41^{***}$ , error rates: *rich* 2 %, *wealthy* 0 %), *quick/rapid*:  $\chi^2 = 1523.19^{***}$ , (*glad/happy*:  $\chi^2 = 3217.41^{***}$ , error rates: *glad* 0 %, *happy* 2 %), (*shut/closed*:  $\chi^2 = 1166.91^{***}$ , error rates: *shut* 65 %, *closed* 5 %).

<sup>27</sup> To account for this astonishing fact, we can offer a hypothesis that is suggested by the normalized frequencies of *shut* and *closed* in our control contexts: While the frequency of *closed* per million words is a

parallel results from analyses of the same adjectives in the 100-million-word *British National Corpus* (BNC).

The fact that the same contrasts hold across diachronically spaced subsections of COHA and written and spoken parts of COCA and BNC (all of which have a differently weighted genre composition) and thus, across both major national varieties, further corroborates the assumption that the distributional contrasts between adjectives may at least in part be attributed to the preference for rhythmic alternation and the spacing of stresses. The fact that this tendency is not necessarily stronger in spoken language than in writing is in line with earlier findings about rhythmic alternation effects in grammatical variation (cf. Schlüter, 2005, p. 291): While the spoken form of words, including their rhythmic properties, is of course the only one relevant in speech, there is substantial evidence that subvocalizations also play an important part in writing and (silent) reading. What is more, speech often takes place under real-time pressures, while writing as a rule involves more intense pre-planning and monitoring, and thus may attain higher levels of well-formedness and optimization of the output: Writers typically invest considerable effort in making their text *sound* good, which no doubt includes its rhythmic dimension (see also the discussion in Schlüter, 2005, p. 50-55). Thus, it comes as no surprise that rhythm turns out to be no less influential in written data than in the spoken parts of COCA.

## 5. Conclusion and perspectives

The analyses described in the present study have produced highly consistent results across speech and writing and across diachronic snapshots of written American English spanning almost two centuries. The results provide substantial cumulative support for our initial hypothesis H<sub>1</sub>, to the effect that adjectives with equivalent meanings and different rhythmic shapes do not occur equally often in all syntactic functions: It has been found that monosyllabic or end-stressed adjectives occur less often before stressed syllables than their non-end-stressed (near) synonyms. The critical contexts for this hypothesis have been pinned down to attributive uses immediately preceding nouns, since nouns in English typically have initial stress. For the minority of nouns that have their primary stress on a later syllable (ca. 15 % of tokens), a carry-over effect has been postulated, such that the avoidance of prenominal use in the large number of rhythmically problematic contexts spills over to unproblematic contexts through a generalized aversion to placement in grammatically equivalent (i.e. prenominal) positions. As a supplementary hypothesis, it has been argued that phonetic duration of a stressed syllable can, under certain circumstances, fulfil the same buffering function as an extra syllable. Some initial evidence has been presented, which suggests that the size of the effect is comparable to that of an extra buffer syllable. In sum, our data indicate that the Principle of Rhythmic Alternation has the potential to influence lexical choices.

While these results are supported by robust quantitative evidence, this is not to deny that there are important limitations to the impact of rhythm: Since our case studies have targeted individual lexemes and each one has idiosyncratic properties, the synonymy presupposed here has proved to be only partial in many respects. Some members of our synonym pairs have larger denotational ranges than others (*rich* vs. *wealthy*, *closed* vs. *shut*), some belong to different stylistic levels and involve slight referential nuances (*fast* vs. *quick* vs. *rapid*) and some lend themselves better than others

---

multiple of that of *shut* in contemporary American English generally, *shut* is still better entrenched in speech than in writing (cf. *Oxford Learner's Dictionaries Online*, s.v. *shut*, *closed*, where it is noted that “[e]specially in *North American English*, **shut** can sound less polite [than **closed**].”). Thus, we propose that the use of *closed* in speech seems to be specifically triggered in rhythmically critical prenominal positions, where as many as 40 % of its occurrences can be found.

to potential (characterizing) or actual (temporary) interpretations, respectively (*happy* vs. *glad*). ‘Fixed collocations’ (phraseological units), on the one hand, can be seen to survive in spite of stress clashes acquired through loss of final schwa (in the case of *glad tidings*), and on the other hand, newly coined units (such as *fast food*, *glad hand* and *glad eye*) may embrace this rhythmically critical condition to indicate pragmatic markedness.

Despite these restrictions, all four case studies have involved lexemes that are morphologically unrelated, but can at least part of the time replace each other, depending on the rhythmic context. These findings contribute important new insights regarding the scope of influence of the Principle of Rhythmic Alternation: The current state of research on the principle has accumulated evidence that it can trigger the avoidance of rhythmically inconvenient morphological or syntactic structures and their replacement by alternative grammatical forms whenever alternatives are available. It is of course easier to assume (near) semantic identity between morphologically marked or unmarked instances of the same lexeme in the same grammatical function, e.g. (17)-(23), or between syntactic constructions containing the same words, e.g. (24)-(29), to name just some of the grammatical phenomena studied in recent literature. (Note, however, that even here specific semantic distinctions have been shown to exist between some of the pairs.)

- (17) *worse* vs. *worser*,
- (18) *drunk* vs. *drunken*, *broke* vs. *broken*, *struck* vs. *stricken*,
- (19) *knit* vs. *knitted*, *lit* vs. *lighted*,
- (20) *quick* vs. *quickly*, *scarce* vs. *scarcely*,
- (21) *go fishing* vs. *go a-fishing*,
- (22) *make someone (to) do something*, *bid someone (to) do something*,
- (23) *dare (to) do something* (all in Schlüter 2005, 2015),
- (24) *give Jim the book* vs. *give the book to Jim* (Anttila, Adams & Speriosu 2010),
- (25) *the teacher’s method* vs. *the method of the teacher* (Ehret, Wolk & Szmrecsanyi 2014, Shih, Grafmiller, Futrell & Bresnan 2015),
- (26) *he knew (that) Lucy washed the dishes* (Lee & Gibbons 2007),
- (27) *green and yellow* vs. *yellow and green* (Mollin 2012, 2014, Lohmann 2014),
- (28) *colleges and universities* vs. *universities and colleges* (Mollin 2012, 2014, Lohmann 2014),
- (29) *John likes beans* vs. *beans, John likes* (Speyer 2010).

Going beyond these fairly well-established facts, we have provided the first pieces of quantitative evidence for Bolinger’s impressionistic observation that the choice of lexical alternatives may equally be determined by their rhythmic properties (1965b, p. 149).

On a theoretical level, the present study supplies a phonologically induced limiting factor for synonym choice within theories of synonymy. More importantly in our context, our study contributes to a more comprehensive assessment of the influence of the Principle of Rhythmic Alternation, arguing that the scope is larger than has so far been shown. Without going into detail here, an assumption shared by many linguistic models and supported by massive and largely undisputed evidence is that language structure as well as language processing are characterized by a hierarchical organization.<sup>28</sup> Thus, language involves a semantic, lexical, syntactic, morphological and phonological level of representation, with semantics at the ‘top’ (second only to pragmatics) and phonology at the ‘bottom’.

---

<sup>28</sup> See, for instance, Bock (1987), Levelt (1989), Lamb (1999), Berg (1988, 1998, 2009). For a review of some models, see Schlüter (2005, p. 237-306) and Lohmann (2014, p. 144-182).

- As has been argued by Schlüter (2015; cf. also 2005, p. 285-291), inside the phonological component, various powerful stress rules are at work that remedy violations of rhythmic alternation (stress deletion, addition and shift).
- Morphological variants such as (17) to (21), being situated close to the phonological level, are numerous and examples of this kind could easily be multiplied, e.g. from among the many variants of weak and strong past participles and suffixed and suffixless adverbs. What is more, these are heavily influenced by rhythmic constellations, the avoidance of stress clashes not infrequently being the prime determinant of their distribution.
- Syntactic structures such as (22) to (29), despite being quite remote from the phonological level, have still been shown to reflect attempts to optimize the rhythm of the outcome. The effect strength is however more limited (cf. Schlüter, 2015; see also Lee & Gibbons, 2007; Shih, Grafmiller, Futrell & Bresnan, 2015; Ehret, Wolk & Szmrecsanyi, 2014) and there are a few studies that failed to produce evidence of stress clash avoidance in multifactorial settings (e.g. Lohmann, 2011, p. 510, 2014, p. 112).
- As for lexical choices, an empirical demonstration of rhythmic influences has been pending up until now – a gap that we hope to have filled with the present study.

We thus embrace a non-modular, though layered model of language and language processing where constraints and influences of one level of representation interpenetrate other levels.<sup>29</sup> Significantly, this influence is not all-pervasive, but decreases as the distance between levels increases.<sup>30</sup> With regard to rhythmic well-formedness constraints, this means that their influence is strongest on the adjacent level of grammatical morphology, moderately strong at the syntactic level and – as has been shown in the present contribution – still demonstrable at the level of lexical selection. While the four case studies selected here yield relatively homogeneous results as to the constancy and size of the rhythmic effects, it has to be admitted that with several other adjective pairs that were tried, POS-tag-based pilot studies did not produce any obvious results. In these cases, no further attempts were made to tease apart potentially conflicting factors which may obscure the expected rhythmic effects. This indicates that the latter may become visible only in the absence of more cogent constraints of other kinds and can be used as a last resort in lexical choices, everything else being equal.

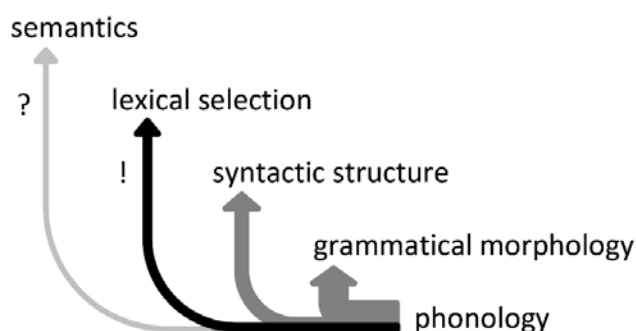


Figure 7: The scope of the Principle of Rhythmic Alternation across levels of linguistic representation

<sup>29</sup> A fully accountable estimation of rhythmic effects on other levels than that of lexical choices is beyond the scope of the present article. For more detailed and comprehensive assessments of the range and strength of rhythm in grammatical morphology and syntax, the reader is referred to the book-length discussion in Schlüter (2005) as well as the overview in Schlüter (2015).

<sup>30</sup> For a congruent model based on a quantitative study of coordinate constructions, see Lohmann (2014, p. 168-173).

Figure 7 presents a rough sketch of the components of a model of language processing.<sup>31</sup> In language production, the major flow of information is from top to bottom.<sup>32</sup> The phonological component is the one that ultimately converts the string of lexemes, syntactic units and morphemes into pronounceable phonological structures. Violations of the Principle of Rhythmic Alternation will cause processing difficulties at this late stage. The width of the arrows symbolizes the size of rhythmic effects, which diminish with increasing distance from the phonological level (where the rhythm factor is located). This can be conceptualized as a feedback mechanism that anticipates violations of rhythmic (and other phonological) principles following from constellations at higher linguistic levels, and works to pre-empt them by promoting competing (grammatical, syntactic or lexical) options. A potential influence of the Principle of Rhythmic Alternation on semantics will have to be explored in future research.<sup>33</sup>

On the diachronic plane, two kinds of hypotheses suggest themselves. First, as the detour into longstanding collocations (i.e. potential phraseological units) with *glad* has shown, the Middle English loss of unstressed final syllables in adjectives is likely to have increased dysrhythmia in sequences of attributive adjectives and nouns (cf. Minkova, 1991, p. 180, quoted above). For many pairs consisting of (now) monosyllabic or end-stressed adjectives and initially stressed nouns, the loss of endings has created loci for stress clashes that can no longer be averted. While in free collocations and newly developing phraseological units such as *happy news*, rhythmically critical adjectives have sometimes come to be replaced by uncritical ones with an unstressed final syllable (as has been argued throughout this paper), individual phraseological units (*fast food*, *glad eye*, *glad hand*) appear to have been coined with deliberate pragmatic marking supported by rhythmic deviancy. Second, we may hypothesize that a syntactic restriction caused by rhythmic difficulties may contribute to a secondary semantic specialization: While *glad(e)* may once have been as common in attributive uses as *happy*, rhythmic inappropriateness may have progressively confined it to predicative uses. This, in turn, may have been a factor in obliterating potential (characterizing) senses of *glad* and promoting actual (temporary) senses, since attributive position is typically associated with potential semantics. In other words, the subtle meaning distinction found in Present-Day English between *glad* and *happy* may in part have arisen from their rhythmic disparity.

The case studies presented in section 4 have different degrees of statistical significance: In some cases, massive ambiguity led to imprecision of the searches (and concomitant statistical insecurity). What is more, some cross-cutting factors (in particular, phraseological fixations) could not be eliminated and may distort the picture. Be that as it may, taken together, all results appear to point in the same direction, in support of our leading hypothesis. In future research, it would thus be useful to study distributions of further adjective pairs with overlapping meanings and rhythmically relevant contrasts. Our search of the “Synonyms” section of *Oxford Learner’s Dictionaries Online* yielded, among others, the following candidates: *afraid – frightened*, *mad – crazy*, *grave – solemn*, *fond – devoted*, *frank – outspoken*, *sour – acid*, *smart – clever*, *superb – excellent*. In view of the sheer size of the English lexicon, the number of (near) synonyms,

<sup>31</sup> For a more detailed elaboration of this model, see Schlüter (2005: 257-306).

<sup>32</sup> In language perception, the direction is reversed, but the focus of this contribution is on the choice of lexical items in the building of utterances.

<sup>33</sup> It has been suggested, for instance, that the insertion of certain buffer elements between clashing stresses may serve primarily rhythmic purposes, but the semantic contribution of these elements – negligible though it may appear – leads to a slight modification of the originally intended communicative contents (cf. Bolinger, 1980, p. 57; Schlüter, 2005, p. 135).



the problems involved in their comparison, and the space allotted to this contribution, we have to leave this work undone for the time being.

## References

### Primary sources

- BNC: Davies, Mark (2004-) *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available online at <http://corpus.byu.edu/bnc/>.
- COCA: Davies, Mark (2008-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>.
- COHA: Davies, Mark (2010-) *The Corpus of Historical American English: 400 million words, 1810-2009*. Available online at <http://corpus.byu.edu/coha/>.
- Cowie, Anthony P.; Mackin, Ronald & McCaig, Isabel R. (1983) *Oxford Dictionary of Current Idiomatic English*. Vol. II: *Phrase, Clause & Sentence Idioms*. Oxford: Oxford University Press. [Repr. 1993 under the title: *Oxford Dictionary of English Idioms*.]
- EEBO: *Early English Books Online* (2003-2015) ProQuest LLC. Available online at <http://eebo.chadwyck.com/home>.
- EPD: Jones, Daniel (2011) *English Pronouncing Dictionary*. Peter Roach, Jane Setter & John Esling (eds.). 18<sup>th</sup> ed. Cambridge: Cambridge University Press.
- OED: *Oxford English Dictionary Online* (2000-) Oxford University Press. Available online at <http://dictionary.oed.com/>.
- Oxford Learner's Dictionaries Online* (2015) Oxford University Press. Available online at <http://www.oxfordlearnersdictionaries.com/>.
- Thesaurus.com (2016) Dictionary.com, LLC. Available online at <http://www.thesaurus.com/>.

### Secondary sources

- Anttila, A., Adams, M. & Speriosu, M. (2010). The role of prosody in the English dative alternation. *Language and Cognitive Processes* 7/8/9, 946-981.
- Baugh, A. C. & Cable, T. (2013). *A History of the English Language*. Sixth edition. London: Routledge.
- Bayer, J. (1989). A note on alleged PRO-government. In D. Jaspers, W. Klooster, Y. Putseys & P. Seuren (Eds.), *Sentential Complementation and the Lexicon: Studies in Honour of Wim de Geest* (pp. 1-20). Dordrecht/Providence: Foris.
- Behaghel, O. (1924). *Deutsche Syntax: Eine geschichtliche Darstellung*. Volume II: *Die Wortklassen und Wortformen*. Heidelberg: Winter.
- Berg, T. (1988). *Die Abbildung des Sprachproduktionsprozesses in einem Aktivationsflußmodell. Untersuchungen an deutschen und englischen Versprechern*. Tübingen: Niemeyer.
- Berg, T. (1998). *Linguistic Structure and Change: An Explanation from Language Processing*. Oxford: Clarendon Press.
- Berg, T. (2009). *Structure in Language: A Dynamic Perspective*. London: Routledge.
- Bock, J. K. (1987). An effect of the accessibility of word forms in sentence structures. *Journal of Memory and Language* 26, 119-137.
- Bolinger, D. L. (1952). Linear modification. *Papers of the Modern Language Association* 67, 1117-1144.
- Bolinger, D. L. (1965a). Binomials and pitch accent. In D. L. Bolinger (Ed.), *Forms of English. Accent, Morpheme, Order* (pp. 129-138). Cambridge, Mass.: Harvard University Press.

- Bolinger, D. L. (1965b). Pitch accent and sentence rhythm. In: D. L. Bolinger (Ed.), *Forms of English: Accent, Morpheme, Order* (pp. 139-180). Cambridge, Mass.: Harvard University Press.
- Bolinger, D. L. (1967). Adjectives in English: Attribution and predication. *Lingua* 18, 1-34.
- Bolinger, D. L. (1980). A not impartial review of a not unimpeachable theory: Some new adventures of ungrammaticality. In: R. W. Shuy & A. Shnukal (Eds.), *Language Use and the Uses of Language* (pp. 53-67). Washington: Georgetown University Press.
- Bolinger, D. L. (1981). *Two Kinds of Vowels, Two Kinds of Rhythm*. Bloomington, In.: Indiana University Linguistics Club.
- Bolinger, D. L. (1986). The English beat: Some notes on rhythm. In G. Nickel & J. Stalker (Eds.), *Problems of Standardization and Linguistic Variation in Present-Day English* (pp. 36-49). Heidelberg: Groos.
- Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. Tübingen: Niemeyer.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cruse, D. A. (2002). Paradigmatic relations of inclusion and identity III: Synonymy. In D. A. Cruse, F. Hundsnurscher, M. Job & P. R. Lutzeier (Eds.), *Lexikologie / Lexicology: Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen / An International Handbook on the Nature and Structure of Words and Vocabularies*. Vol. I (pp. 485-497). Berlin: Mouton de Gruyter.
- Cutler, A. & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2, 133-142.
- Dobson, E. J. (1968). *English Pronunciation 1500-1700*. Volume II: *Phonology*. Second edition. Oxford: Clarendon Press.
- Dolezal, F. (2013). Synonymy and sameness of meaning: An introductory note. *International Journal of Lexicography* 26(3), 255-259.
- Ehret, K., Wolk, C. & Szmrecsanyi, B. (2014). Quirky quadratures: On rhythm and weight as constraints on genitive variation in an unconventional data set. *English Language and Linguistics* 18(2), 263-303.
- Fijn van Draat, P. (1910). *Rhythm in English Prose*. Amsterdam: Swets & Zeitlinger N.V. [Repr. 1967.]
- Gray, B. & Biber, D. (2015). Phraseology. In D. Biber (Ed.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 125-145). Cambridge: Cambridge University Press.
- Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry* 15, 33-74.
- Jacobsson, B. (1996). A new look at 'predicative-only' adjectives in English. *Journal of English Linguistics* 24, 206-219.
- Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in Language* (pp. 350-377). Cambridge, Mass.: MIT Press; New York & London: J. Wiley & Sons.
- Jespersen, O. (1909). *A Modern English Grammar on Historical Principles*. Part I: *Sounds and Spellings*. London: George Allen & Unwin/Copenhagen: Munksgaard.
- Kager, R. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Ph.D. diss. Rijksuniversiteit Utrecht.
- Knappe, G. (2012). Idioms and fixed expressions. In L. Brinton & A. Bergs (Eds.), *English Historical Linguistics: An International Handbook*. Vol. I (pp. 177-196). Berlin: Mouton de Gruyter.
- Lamb, S. M. (1999). *Pathways of the Brain: The Neurocognitive Basis of Language*. Amsterdam/Philadelphia: Benjamins.
- Lee, M.-W. & Gibbons, J. (2007). Rhythmic alternation and the optional complementiser in English: New evidence of phonological influence on grammatical encoding. *Cognition* 105, 446-456.

- Leisi, E. (1985). *Praxis der englischen Semantik*. Second edition. Heidelberg: Winter.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, Mass./London, UK: MIT Press.
- Lohmann, A. (2011). *Help vs help to*: A multifactorial, mixed-effects account of infinitive marker omission. *English Language and Linguistics* 15(3), 499-521.
- Lohmann, A. (2014). *English Coordinate Constructions: A Processing Perspective on Constituent Order*. Cambridge: Cambridge University Press.
- Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press. [Reprinted 1989.]
- Minkova, D. (1991). *The History of Final Vowels in English: The Sound of Muting*. Berlin/New York: Mouton de Gruyter.
- Mollin, S. (2012). Revisiting binomial order in English: Ordering constraints and reversibility. *English Language and Linguistics* 16(1), 81-103.
- Mollin, S. (2014). *The (Ir)reversibility of English Binomials*. Amsterdam: Benjamins.
- Naciscione, A. (2010). *Stylistic Use of Phraseological Units in Discourse*. Amsterdam/Philadelphia: John Benjamins.
- Nespor, M. & Vogel, I. (1989). On clashes and lapses. *Phonology* 6, 69-116.
- Paradis, C. (1997). *Degree Modifiers of Adjectives in Spoken British English*. Lund: Lund University Press.
- Roth, T. (2014). *Wortverbindungen und Verbindungen von Wörtern: Lexikografische und distributionelle Aspekte kombinatorischer Begriffsbildung zwischen Syntax und Morphologie*. Tübingen: Francke.
- Schlüter, J. (2005). *Rhythmic Grammar: The Influence of Rhythm on Grammatical Variation and Change in English (TiEL 46)*. Berlin/New York: Mouton de Gruyter.
- Schlüter, J. (2008). Constraints on the attributive use of 'predicative-only' adjectives: A reassessment. In N. Gisborne & G. Trousdale (Eds.), *Constructional Explanations in English Grammar (TiEL)* (pp. 164-179). Berlin/New York: Mouton de Gruyter.
- Schlüter, J. (2015). Rhythmic influence on grammar: Scope and limitations. In R. Vogel & R. van de Vijver (Eds.), *Rhythm in Cognition and Grammar: A Germanic Perspective* (pp. 179-205). Berlin/Munich/Boston: De Gruyter Mouton.
- Selkirk, E. O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass./London, UK: MIT Press.
- Shih, S., Grafmiller, J., Futrell, R. & Bresnan, J. (2015). Rhythm's role in genitive construction choice in spoken English. In R. Voge & R. van de Vijver (Eds.), *Rhythm in Cognition and Grammar: A Germanic Perspective* (pp. 207-233). Berlin/Munich/Boston: De Gruyter/Mouton.
- Speyer, A. (2010). *Topicalization and Stress Clash Avoidance in the History of English*. Berlin/New York: De Gruyter Mouton.
- Štekauer, P. (2000). *English Word-formation: A History of Research (1960-1995)*. Tübingen: Narr.
- Xiao, R. (2015). Phraseology. In D. Biber (Ed.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 106-124). Cambridge: Cambridge University Press. 106-124.