

English Language and Linguistics, 26.4: 807–831. © The Authors, 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

doi:[10.1017/S1360674322000065](https://doi.org/10.1017/S1360674322000065)

Phrasal verbs in Early Modern English spoken language: a colloquialization conspiracy?¹

PAULA RODRÍGUEZ-PUENTE  and MARÍA OBAYA-CUELI
University of Oviedo

(Received 1 October 2021; revised 31 March 2022)

Phrasal verbs (e.g. *fade away*, *give up*) tend to be associated with spoken, colloquial registers, not only in Present-day English, but also in previous stages of the language. This view has recently been challenged by Thim's (2006a, 2012) 'colloquialization conspiracy', according to which the idea that phrasal verbs are colloquial is based on a misconception which first arose in the eighteenth century. In the current study we seek to verify Thim's claim by exploring phrasal verbs in *A Corpus of English Dialogues 1560–1760*, a 1.2-million-word corpus of Early Modern English (EModE) speech-related text types. Based on a sample of over 7,000 examples, we demonstrate that the linguistic features, distribution and high productivity of phrasal verbs in the EModE period point towards a full entrenchment of these combinations in the spoken language, which leads us to the conclusion that the colloquial status of phrasal verbs in EModE is not merely a matter of a 'colloquialization conspiracy'.

Keywords: phrasal verbs, Early Modern English, colloquialization conspiracy, spoken language of the past, corpus analysis

1 Introduction

Phrasal verbs (e.g. *fade away*, *give up*) tend to be associated with spoken colloquial registers, not only in contemporary English (see, e.g., Cowie & Mackin 1975: iv; McArthur 1989: 40; Biber *et al.* 1999: 408, 409; Liu 2011: 675), but also in previous stages of the language (see, e.g., Hiltunen 1994; Claridge 2000: 185–97; Kytö & Smitterberg 2006; Smitterberg 2008; Rodríguez-Puente 2017, 2019). As far as Present-day English (PDE) is concerned, Dempsey *et al.* (2007) used phrasal verbs to distinguish computationally between spoken and written registers, concluding that these combinations can be indicative of the degree of spokenness and formality of a given text. Likewise, phrasal verbs are a linguistic feature with a positive load on Dimension 1 'Oral vs. Literate Discourse' in Biber's (2003) multi-dimensional analysis of variation between university spoken and written registers. As regards earlier periods,

¹ For generous financial support, we are grateful to the Spanish Ministry of Science and Innovation (grant PID2020-114604GB-I00). We would like to thank the two anonymous reviewers and the journal editors, especially Laurel J. Brinton, for helpful comments and feedback.

in a corpus-based study of phrasal verbs from 1650 to 1999, Rodríguez-Puente (2019) analyzed these combinations in ten different registers, demonstrating that ‘phrasal verbs significantly distinguish spoken from written registers and informal from formal registers’ (2019: 282). For her, the high occurrence of phrasal verbs in the trial proceedings of the *Old Bailey Corpus*, a register which closely represents spoken language and also lower social strata of society (Huber 2007), is a clear indication that phrasal verbs were a feature of the spoken language in the Late Modern English (LModE) period (2019: 264–77). However, Rodríguez-Puente (2019) observes that the subject matter and the idiosyncrasies and particular style of writers can also condition the appearance of phrasal verbs, with certain combinations being particularly useful to describe events in the narrative sections (see also Hiltunen 1994: 138).

With reference to the Early Modern English (EModE) period, however, the view that phrasal verbs are typical of spoken, colloquial registers has been challenged by Thim (2006a, 2012: 233–46), who argues that such an idea is based on a ‘colloquialization conspiracy’, a misconception which took off in the eighteenth century. He claims that the recurrent assumption that EModE phrasal verbs are typical of colloquial, everyday registers is based on the common view that PDE phrasal verbs are colloquial. Thim (2006a) examines a sample of over 2,000 phrasal verbs extracted from *Everyday English* (Cusack 1998), a collection of 64 non-literary EModE texts which include the speech production of members of the lower social classes between 1500 and 1684. However, the author does recognize that, while some of the texts in this collection portray renderings of the spoken language, ‘others seem conceptionally literate’ (Thim 2006a: 292), so that the assignment of levels of orality is dubious in some cases. Nonetheless, from the analysis of phrasal verbs in this collection, Thim concludes that EModE phrasal verbs are ‘stylistically neutral’ (2006a: 302). For him, the use of these combinations is motivated by the subject matter of the text, rather than the medium (spoken or written) or the degree of formality (see also Thim 2011). He subsequently compared the frequencies of phrasal verbs provided in several diachronic and contemporary studies (2012: 211–14), questioning their reliability and arguing that most statements made so far about the quantitative diachronic development of phrasal verbs throughout the EModE period are ‘meaningless’ (2012: 213), as it is not possible to find any clear trends in the historical development of these constructions. For him, the colloquiality of phrasal verbs has been taken for granted in studies such as Hiltunen (1994), Claridge (2000) and Blake (2002), ‘even if the evidence does not support that assumption’ (Thim 2012: 216; see further Thim 2006a: 298–9).

Despite his criticism of previous research, Thim (2012: 213) recognizes that the analysis of a larger diachronic corpus to trace the development of phrasal verbs is necessary to arrive at new long-term comparisons, a task which he describes as ‘inadvisable in an untagged corpus’ (2012: 213), in that a great deal of manual analysis would be necessary to discern phrasal particles from their homonymous prepositions. With this article we have assumed such an inadvisable task in an attempt to verify Thim’s (2006a, 2012) argument by exploring phrasal verbs in *A Corpus of English Dialogues 1560–1760* (CED; Kytö & Culpeper 2006). CED contains EModE

speech-related texts in which constructed and real dialogues are represented (see section 3), thus providing a more reliable source of speech-like production of the period than Cusack's (1998) collection. Our goal is to analyze the most salient linguistic features of phrasal verbs, their frequency, usage and productivity in a large sample of EModE speech-related texts, including a wide number of particles and a transparent semantic classification along the lines of Rodríguez-Puente (2019: 75–84). In light of our findings, we also revisit previous studies criticized by Thim in order to provide an alternative explanation for those apparently chaotic paths of distribution which he finds in them. We thus contribute to the existing literature with a robust quantitative diachronic study that allows us to describe phrasal verbs empirically and objectively in the EModE period and to attain solid conclusions about their alleged colloquial status at that time.

The article is structured as follows. Section 2 provides an overview of phrasal verbs and their historical development, paying special attention to their status in the EModE period. Section 3 describes the corpus and the methodology used. In section 4 we present our main results and the analysis of our findings, focusing first on the most salient linguistic aspects of the combinations (section 4.1), and then moving on to an analysis of their frequencies and register distribution (section 4.2). Section 5 closes the article with some conclusions.

2 On English phrasal verbs

Phrasal verbs in this article are understood as

(partially) lexicalised and idiomatised two-word lexical units made up of a verb and a particle of adverbial origin, whose syntactic bondedness can be described in terms of a gradient from lowly lexicalised to fully lexicalised and whose meaning can be analysed along a scale from fully transparent to fully idiomatised. (Rodríguez-Puente 2019: 151)

Phrasal verbs are thus distinguished from other related categories, such as prepositional verbs (*look into* in (1)), where the particle is a preposition, and phrasal-prepositional verbs (*look up to* in (2)), which contain both an adverb and a preposition.

- (1) They are **looking into** the matter very carefully.
- (2) Most people **look up to** celebrities.

Although the set of possible particles which can be said to be used for the creation of phrasal verbs varies from one author to another, for our purposes and for the sake of comparison with previous work, we have included those listed by Rodríguez-Puente (2019: 44), which are based on compilations provided in previous studies (see, among others, Bolinger 1971; Cowie & Mackin 1975; Fraser 1976; Quirk *et al.* 1985: 1151; Hiltunen 1994; Claridge 2000: 46). This list includes the following thirty-four items:

aback, aboard, about, above, across, after, ahead, along, apart, around, aside, astray, asunder, away, back, behind, by, counter, down, forth, forward(s), home, in, off, on, out, over, past, round, through, to, together, under, up

Whereas a set of possible particles is easy enough to establish, verbs are a different matter. In principle, any kind of lexical verb can function as the verbal element in a phrasal

combination, although they are most commonly (but not necessarily) of native origin and generally ‘monosyllabic or disyllabic verbs with the accent on the first syllable’ (Claridge 2000: 54; see also Martin 1990: 115; Thim 2006a: 219).

Despite having relatively clearly identifiable parts, phrasal verbs are a fuzzy category and constitute an example of how categories can be graded, with some combinations being more prototypical than others in terms of their semantic and syntactic properties (see Rodríguez-Puente 2019: 107–52 for discussion). Following traditional classifications,² we include literal (non-idiomatic) combinations (e.g. *bring up* in (3)), metaphorical or figurative combinations, whose meaning is quite transparent but somehow removed from its original connotation (e.g. *bring up* in (4)), and idiomatic or non-compositional categories, whose interpretation cannot be deduced from the meaning of the parts (e.g. *bring up* in (5)). Along the lines of Rodríguez-Puente (2019: 55–84), we further distinguish aspectual or aktionsart combinations (e.g. *eat up* in (6)), where the particle contributes to the compound with an aspectual or aktionsart meaning (see further Denison 1985; Brinton 1988), reiterative combinations (e.g. *rise up* in (7)), in which the particle repeats part of the meaning already conveyed by the verb, and emphatic combinations (e.g. *barrel up* in (8)), where the meaning of the verb is exactly the same as when appearing alone, but the particle contributes a more colloquial tone and facilitates the division of labor between the verb and the particle, thus having an effect on the information structure of the clause (for further discussion, see Rodríguez-Puente 2013, 2019: 61–70).³

- (3) They **brought** the wine **up** from the cellar.
- (4) They **brought up** the player from the minor leagues.
- (5) They **brought up** an interesting topic in our conversation this morning.
- (6) Termites have **eaten up** the wooden fence.
- (7) He **rose up** and left.
- (8) Let it stand by till cold, and then **barrel it up** for use.

Syntactically speaking, phrasal verbs can be transitive or intransitive. When transitive, the particle can appear at either side of the object when it is a noun phrase (*bring [up] the wine [up]*), whereas it must follow the object when it is a (non-stressed) pronoun (*bring it up* vs **bring up it*).

Phrasal verbs have been studied widely from a diachronic perspective, in terms of their development, their particular semantic and morphosyntactic features, as well as their alleged association with the spoken, colloquial language.⁴ Although these combinations were already attested in OE, it is generally agreed that their frequency

² See Bolinger (1971), Makkai (1972), Denison (1981), Claridge (2000), Thim (2006a, 2012: 11–20), among others.

³ Note that in the *Oxford English Dictionary* (OED), the definition of the verb *barrel* is given under 1(a) as ‘put, pack, store up, or stow away, in a barrel or barrels’, whereas in 1(b) it is clarified that the verb is ‘often emphasized with *up*’ with no change in meaning.

⁴ For a comprehensive list of publications dealing with phrasal verbs in different periods of English, see Rodríguez-Puente (2019: 13n) and references therein.

(in terms of types and tokens) has increased greatly over time, reflecting the analytic drift which has affected English historically (see Spasov 1966: 18–22; Denison 1998: 223; Hiltunen 1999: 133; Rodríguez-Puente 2019: 175–7). The presence of phrasal verbs in modern English is now a common feature in all registers (see Claridge 2000: 104), although they are more commonly found in colloquial, informal language than in formal, written registers (see, among others, Dempsey *et al.* 2007).

By the EModE period the phrasal verb was fully established in the language (Kennedy 1920: 13–14; Konishi 1958: 121; Brinton 1988: 187; Blake 2002: 25). Syntactically, EModE phrasal verbs seem to behave very much like their PDE counterparts (Hiltunen 1994: 133; Claridge 2000: 98), although, according to Castillo (1994: 450), Shakespearean English is more permissive than PDE grammar in allowing alternations of the standard order of constituents.

As regards semantics, it seems that the majority of the EModE phrasal-verb combinations tend to be concrete or literal, ‘with only incipient metaphorical developments in certain contexts’ (Hiltunen 1994: 132). According to Thim (2006a: 296, 2006b: 224–5), non-compositional phrasal verbs are rare before 1600 and their degree of opacity may vary. Therefore, although during the EModE period there is an increase of metaphorical and idiomatic meanings, as well as of aspectual/aktionsart uses (Konishi 1958: 122; Claridge 2000: 96), these are probably not as abundant as in PDE.

3 Corpus and methodology

The features of EModE phrasal verbs described in section 2 are examined in the present article in light of new data extracted from CED.

CED is a 1.2-million-word corpus of EModE speech-related text types running from 1560 to 1760. Although the texts in CED are recorded in the written medium, they contain dialogic exchanges which can be used to explore representations of the spoken language from the past. These can be divided into two broad categories (see Kytö & Walker 2006: 12): authentic dialogue, that is, written records of real speech events (trial proceedings and witness depositions), and constructed or fictional dialogue (drama comedy, handbooks and prose fiction). The group of handbooks consists of instructional or informational texts presented in the form of dialogues between a master and a student, and is further divided into two subcategories, ‘Language Teaching’ and ‘Other’. CED includes an additional subgroup of miscellaneous dialogues of various kinds (Kytö & Walker 2006: 12, 24), but given their mixed nature, they have been omitted from the current analysis.⁵ Table 1 summarizes the overall structure of the corpus texts used in this study.

For a more fine-grained analysis of the results, we divide the texts in CED following Culpeper & Kytö’s (2010: 17–18) classification of speech-related texts into speech-like (e.g. personal letters and diaries), speech-based (e.g. trial proceedings and witness depositions) and speech-purposed (e.g. dramatic dialogues and sermons).

⁵ For a similar approach, see Culpeper & Kytö (2010: 21–60).

Table 1. *Overall structure of CED*

Period	Trial proceedings	Witness depositions	Drama comedy	Didactic works: Language teaching	Didactic works: Other	Prose fiction	Totals
1560–99	19,940	42,080	47,590	16,410	24,750	39,380	190,150
1600–39	14,430	39,930	47,700	17,620	39,370	43,460	202,510
1640–79	70,010	46,820	47,590	16,540	16,310	49,290	246,560
1680–1719	96,630	26,500	47,200	19,790	58,280	47,360	295,760
1720–60	84,650	17,610	48,510	4,030	23,540	44,400	222,740
Total	285,660	172,940	238,590	74,390	162,250	223,890	1,157,720

The examples of phrasal verbs were retrieved automatically from the corpus using *WordSmith Tools* version 8 (Scott 2020). Since CED is not morphologically tagged, the procedure required searching for the individual particles by means of concordances which included all their possible spellings in the EModE period.⁶ This was followed by further manual analysis to identify those cases in which the particles were part of a phrasal verb and to discard homonymous prepositions (for a similar approach, see, among others, Rodríguez-Puente 2019: 44–5).

Differences in the number of words between the various sections of the corpus were considered for the analysis by normalizing the raw data when necessary. We further verified the statistical significance of our results by applying the Kruskal–Wallis test and the Wilcoxon test⁷ by means of the free software R version 4.0.3 (R Core Team 2020). As usual in linguistic analyses, the threshold for statistical significance was set at $p < 0.05$.

4 Results: phrasal verbs in CED

Corpus searches yielded 7,289 examples, a volume of occurrence which indicates that phrasal verbs are well represented in the spoken language of the EModE period, and which is sufficient to describe developmental trends. In this section we first address the most salient linguistic aspects of the combinations, and then move on to an analysis of their frequencies and register differences.

4.1 *Linguistic aspects of phrasal verbs in CED*

The corpus provided examples for most of the particles included in the analysis (see section 2). Five of them, however, were not attested in the sample texts (*above*, *around*,

⁶ All the possible spellings of the particles were identified by means of the WordList feature of *WordSmith*, as well as by examining their corresponding entries in the *OED*.

⁷ The Kruskal–Wallis rank sum test is a non-parametric method used for comparing more than two samples that are independent or not related. When the Kruskal–Wallis test leads to significant results, then at least one of the samples is different from the other samples. In such case, a non-parametric test, such as the Wilcoxon test, can be applied for a more fine-grained analysis (see further Brezina 2018: 195ff.).

Table 2. *Diachronic distribution of phrasal verb particles in CED*
(raw and normalized frequencies per 100,000 words)

	1560–99	1600–39	1640–79	1680–1719	1720–60	Total
	Raw/NF	Raw/NF	Raw/NF	Raw/NF	Raw/NF	Raw/NF
<i>aback</i>	–	–	1/0.40	–	–	1/0.08
<i>aboard</i>	–	1/0.49	1/0.40	1/0.33	1/0.44	4/0.34
<i>about</i>	17/8.94	8/3.95	14/5.67	31/10.48	15/6.73	85/7.34
<i>across</i>	–	–	–	–	3/1.34	3/0.25
<i>after</i>	2/1.05	2/0.98	1/0.40	3/1.01	–	8/0.69
<i>ahead</i>	–	–	–	–	5/2.24	5/0.43
<i>along</i>	5/2.62	10/4.93	34/13.78	41/13.86	14/6.28	104/8.98
<i>apart</i>	2/1.05	1/0.49	–	–	1/0.44	4/0.34
<i>aside</i>	8/4.20	12/5.92	15/6.08	18/6.08	8/3.59	61/5.26
<i>asunder</i>	1/0.52	–	–	2/0.67	–	3/0.25
<i>away</i>	165/86.77	106/52.34	147/59.62	194/65.59	192/86.19	804/69.44
<i>back</i>	27/14.19	28/13.82	61/24.74	61/20.62	56/25.14	233/20.12
<i>behind</i>	1/0.52	5/2.46	3/1.21	7/2.36	4/1.79	20/1.72
<i>by</i>	16/8.41	20/9.87	20/8.11	16/5.40	14/6.28	86/7.42
<i>down</i>	110/57.84	121/59.75	128/51.91	170/57.47	263/118.07	792/68.41
<i>forth</i>	25/13.14	37/18.27	41/16.62	11/3.71	13/5.83	127/10.96
<i>forward(s)</i>	13/6.83	12/5.92	10/4.05	7/2.36	7/3.14	49/4.23
<i>home</i>	77/40.49	82/40.49	50/20.27	84/28.40	57/25.59	350/30.23
<i>in</i>	106/55.74	110/54.31	165/66.92	194/65.59	149/66.89	724/62.53
<i>off</i>	25/13.14	40/19.75	76/30.82	120/40.57	117/52.52	378/32.65
<i>on</i>	46/24.19	60/29.62	86/34.87	162/54.77	82/36.81	436/37.66
<i>out</i>	224/117.8	174/85.92	258/104.63	335/113.26	309/138.72	1,300/112.28
<i>over</i>	21/11.04	16/7.90	42/17.03	35/11.83	37/16.61	151/13.04
<i>round</i>	1/0.52	–	1/0.40	6/2.02	8/3.59	16/1.38
<i>through</i>	1/0.52	4/1.97	7/2.83	8/2.70	8/3.59	28/2.41
<i>to</i>	2/1.05	1/0.49	1/0.40	–	6/2.69	10/0.86
<i>together</i>	19/9.99	25/12.34	27/10.95	21/7.10	10/4.48	102/8.81
<i>under</i>	–	–	1/0.40	2/0.67	–	3/0.25
<i>up</i>	162/85.19	226/111.59	285/115.59	352/119.01	377/169.25	1,402/121.1

astray, *counter* and *past*). As shown in table 2, the least common particles are *aback*, *across*, *asunder*, *under*, *aboard*, *apart* and *ahead*; conversely, *up*, *out*, *away*, *down*, *in*, *on* and *off* are the most frequently used particles in the corpus.

As shown in table 2, a rising trend is observed in most particles, which indicates that phrasal combinations were growing in number during the period. One notable exception is the particle *forth* which reduces considerably in frequency (see also Hiltunen 1994: 134; Martin 1990: 111; Nevalainen 1999: 423; Los 2004: 98; Akimoto 2006; Ishizaki 2009: 45–50; De Smet 2010; Rodríguez-Puente 2019: 167). In our eighteenth-century data *forth* is only attested in trial proceedings, always in the quite fixed and formulaic combination *set forth*, with reference to a deposition or an indictment, as in (9).

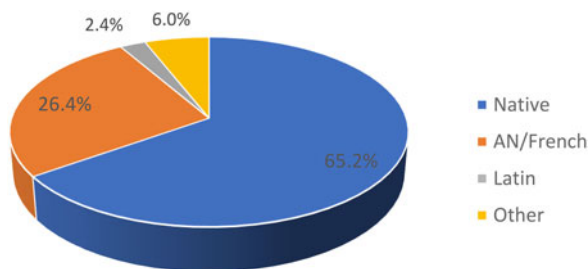


Figure 1. Distribution of the origin of the verbal elements of the phrasal verbs in CED

- (9) Then his Deposition was read, **setting forth**, that he was first Lieutenant of the Royal-Oak the 11th of February 1743. (*Court-Martial ... Captain John Ambrose*, 1745)

Whereas the set of possible particles is rather limited, the number of verbs which can enter into these combinations in EModE is much wider. In our data there are in fact 420 verbal bases. For a more fine-grained analysis we considered the etymology of the verbs using the *OED*. As expected, a significant majority of the verbal bases are native (65.2 percent), although verbs of French and Anglo-Norman (AN) origin are also well represented (26.4 percent), as shown in figure 1.⁸ Verbs of Latin origin are scarce, which comes as no surprise considering the dialogic nature of the texts in the corpus. Thim (2006b: 219) also finds very low rates of French verbs (5 percent) and no examples of Latin bases in the fifteenth- and sixteenth-century letters of the *Corpus of Early English Correspondence* (CEEC). Our proportion of Latin verbs is, in fact, half the size of that found by Rodríguez-Puente (2019: 159), whose study also included formal, written texts.

The verbal bases are primarily monosyllabic or disyllabic verbs with the accent on the first syllable. However, twenty-four of our examples do not follow this general tendency. Five verbs contain three syllables and nineteen are disyllabic with the accent on the second syllable. Combinations containing these verbs are scarce and, in most cases, attested only once. Curiously enough, most of them are cases of reiterative combinations, such as *return back* in (10), and emphatic combinations, such as those illustrated in (11) to (13).

- (10) Did then alsoe see the boy Jackson [...] **returned back** to his master. (*High Commission Court... Diocese of Durham*, 1627–37)
- (11) I will make haste with my Copies, that I may **dispatch them away** this night. (*Trial of Edward Coleman*, 1678)
- (12) Are you sure they [the letters] were never **deliver'd out**? (*Trial of Francis Francia*, 1716)
- (13) [...] I think it was by Mr. Giffard's Direction who **subpoena'd me up**. (*Trial of James Annesley and Joseph Redding*, 1742)

⁸ The category 'Other' includes onomatopoeic and imitative verbs, verbs of unknown origin and one verb each from Spanish, Italian and Dutch.

As argued by Rodríguez-Puente (2019: 61–70; see also Rodríguez-Puente 2013), emphatic combinations are those which contain an apparently superfluous particle, in that the meaning of the compound is not different from that of the simple verb. However, the addition of such particles helps to render the verb ‘more colloquial, informal, or familiar in tone and to make it more salient in the clause’ (2019: 62). Reiterative combinations, in turn, contain a particle that repeats the meaning denoted by the verb. In both types of combinations, the particle is somehow unnecessary. It seems, therefore, that speakers spontaneously add those particles to polysyllabic verbs of foreign origin (and other verbs) on analogy with native combinations as a means of rendering them more colloquial and familiar. The prolific attestation of these combinations in the data lends evidence to the idea of the colloquial character of the texts in CED.

Among these combinations, one curious case is *subpoena up* (see (13)). The combination is produced by a court witness, Samuel Sylvester. The use of the verb *subpoena* by a witness is in itself surprising, as this is a specialized word typical of legal language. The fact that he uses it with the particle *up* may be an indication that he is not completely satisfied with such a formal, specialized word, and that he feels the need to add a colloquializing element, such as *up*.

Having described the two members of the compound separately, we can now move on to the analysis of the combinations. A total of 992 different phrasal verbs were found in CED, including literal combinations but also combinations whose meanings cannot be deduced from those of their parts. *Come in* appears only in a literal sense in the corpus, whereas *fall out* and *find out* are only used idiomatically. Example (14) contains two different uses of *fall out*, the first meaning ‘argue’ and the second ‘happen’. In turn, *come up*, for example, can be used literally (15), metaphorically (16) and with various idiomatic meanings, as in (17), where it means ‘appear’.

- (14) [...] afterward she **fell out** with that man, and vpon this his horse died, she confessed she sent the spirit, how could all things **fall out** so fit? (*Dialogue Concerning Witches*, 1593)
- (15) As it was difficult **coming up**, so (so far as I can see) it is dangerous going down. (*Pilgrim's Progress*, 1678)
- (16) All, all, good Will or to stay thy fury till my Rents **come up**. (*The Wild Gallant*, 1669)
- (17) A head **comes vp** with eares of Come, and she combs them in her lap. (*The Old Wiues Tale*, 1595)

Providing a full analysis of the semantic types of phrasal verbs goes beyond the scope of this article. However, our data indicate that metaphorical and idiomatic meanings are well represented in CED, which may be accounted for in terms of the colloquial nature of the texts. Hiltunen (1994), who analyzed eight different registers from the EModE section of the *Helsinki Corpus* (HC), finds that most collocations there are concrete (1994: 132; see also Thim 2006a: 296, 2006b: 224–5). He explains that phrasal verbs are less frequent in written than in colloquial language because ‘metaphorical combinations tend to be even more marked as colloquial and therefore avoided in the written language’ (1994: 139n). In our data, metaphorical and figurative combinations are found as early as the sixteenth century, which is, therefore, another indicator of the

colloquial character of phrasal verbs in the period. Besides *fall out* in (14), other non-compositional combinations found in CED in subperiod 1560–1599 include *bring up* ‘rear’, *face down* ‘dispute’, *find out* ‘discover’, *give down* ‘let flow milk’, *give in* ‘admit’, *set down* ‘put down in writing’, *take up* ‘occupy’ and *turn to* ‘apply oneself to some task or occupation’. Some illustrative examples are provided in (18) to (20).⁹

(18) In Thessalie was I borne and **brought vp**. (*The Old Wiues Tale*, 1595)

(19) [...] his sayde beasts did **giue downe** blood in steede of milke [...]. (*Witches, Taken at S. Oses*, 1582)

(20) Is your husband **turned to**? (*Dialogve Concerning Witches*, 1593)

Another indication that phrasal verbs are well entrenched in the spoken language of the EModE period is their occurrence in several idiomatic expressions, such as *fall together by the ears* (‘be at variance, fall out’ (*OED* s.v. *by the ear* c.(d.) in *ear* n.)), as in (21), *give in verdict* and *take up arms*.

(21) [...] in short time they two **fell together by the eare**. (*A Nest of Ninnies*, 1608)

Among the combinations attested in CED we also found a great number of *hapax legomena* and *dislegomena*, which, as demonstrated by Baayen (1989, 1992, 1993), constitute a good measure of the productivity of a given construction. Applying such a measure, phrasal verbs seem to be highly productive in CED, in that 456 out of 992 combinations are used only once in the corpus, whereas 149 are attested only twice. Many of the single-occurrence items are, however, nonce formations no longer found in PDE (and not even recorded in the *OED*), such as some of the examples with polysyllabic verbs mentioned above, which seem to be spontaneous creations typical of spoken interaction. Curious examples include *weigh up* (22) and *know through* (23), unusual not only for being attested only once in the corpus but also because they are formed with stative verbs, usually described in the literature as rare in the creation of phrasal verbs.

(22) And then he would be a fort’night **weighing vp** againe. (*Bartholmew Fayre*, 1631)

(23) [...] were they large enough to see, and **know** a Person **through**? (*Tryal ... of John Stevenson*, 1759)

4.2 Diachronic distribution and cross-register analysis

In this section we examine the diachronic and register distribution of phrasal verbs in CED. Figure 2 illustrates their diachronic development across the various subperiods represented in the corpus.¹⁰

⁹ Note that example (19) antedates the first recorded example in the *OED*, which is from 1700 (*OED* s.v. *give down* in *give* v.). Similarly, (20) is a much earlier example than the first recorded example in the *OED*. Under the entry *turn to* (*OED* s.v. *turn to* in *turn* v.) the first quotation dates from 1814. The dictionary provides a cross-reference to *turn* 28c for a related meaning, dated 1667.

¹⁰ Raw frequencies are as follows: 1560–1599: 1,076; 1600–1639: 1,101; 1640–1679: 1,475; 1680–1719: 1,881; 1720–1760: 1756.

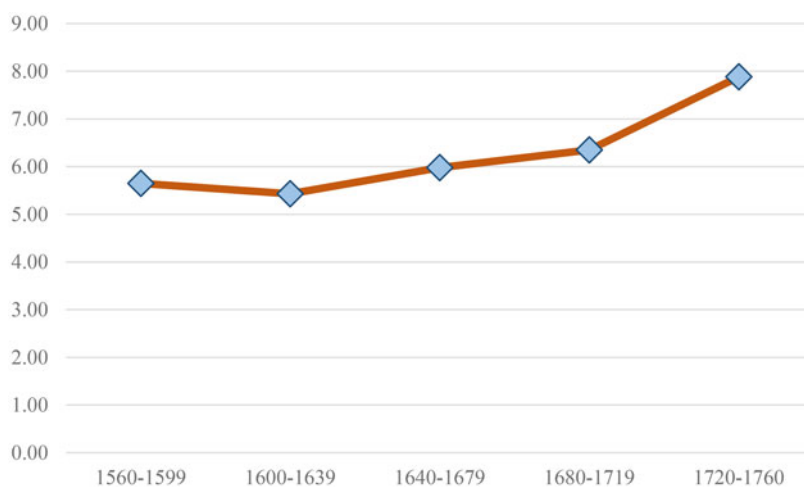


Figure 2. Diachronic distribution of tokens of phrasal verbs in CED (normalized frequencies per 1,000 words)

As can be seen, although there is a slight decrease between the first and second subperiods, phrasal verbs grow significantly from 1600–39 to the second half of the eighteenth century ($p = 0.02598$), a rise which is particularly pronounced between the last two subperiods.

Previous studies indicate that phrasal verbs begin to increase in frequency from 1700 onwards (see Konishi 1958: 125; Spasov 1966: 125; Martin 1990; Wild 2010: 227), although this depends to a large extent on the type of texts examined (see Rodríguez-Puente 2019: 175–7). In CED the sharpest increase is recorded between the last two subperiods, yet phrasal verbs experience a slow but general growth from 1600 onwards, much earlier than usually reported in the literature, which may be considered another indication of their consolidated status in the speech-like texts of the period. Previous studies also point towards a downturn in the growth of phrasal verbs during the eighteenth and the first half of the nineteenth centuries, probably aided by the prescriptivist ideas of the time, which encouraged the use of Latin forms and criticized particles for being superfluous and for their connection with monosyllables and stranded prepositions (see especially Yáñez-Bouza 2015). Such a reversal is, however, not attested in CED, perhaps because of the dialogic nature of the texts involved or simply because we do not have data beyond 1760 and the change is yet to be seen in spoken registers. However, a closer look at the evolution of individual registers is necessary to confirm this general trend (see below for further discussion).

The data presented in figure 2 indicate that the total frequency of combinations is higher at the end of the period than at the beginning. Yet a look at the frequencies of types (figure 3) shows that this is not accompanied by an increase in their type productivity. However, it must be borne in mind that these data reflect the total number of combinations, but not their different (literal and/or idiomatic) connotations. As

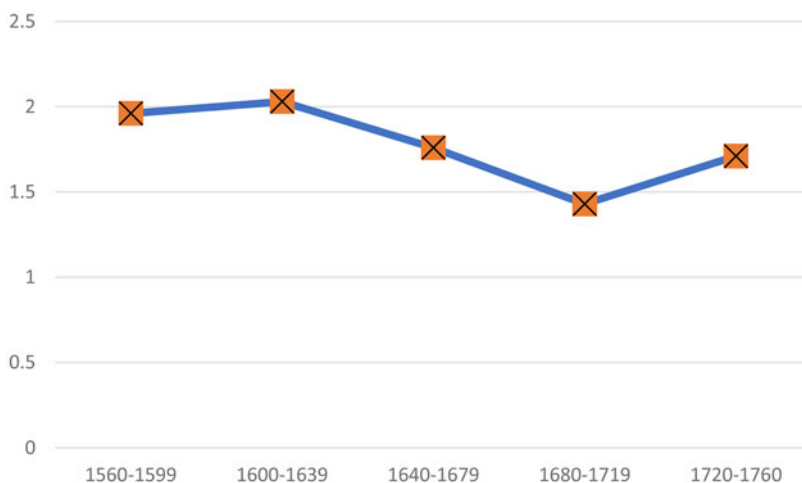


Figure 3. Diachronic distribution of types of phrasal verbs in CED, 1560–1760 (normalized frequencies per 1,000 words)

shown in (3) to (5), one single combination may have many varied meanings, which are normally literal when they are first introduced into the language and later idiomatic (see especially Rodríguez-Puente 2012, 2019: 123–38; Brown & Palmer 2015).

Moreover, the data represented in figures 2 and 3 refer to the whole corpus. Although standard English grammar and vocabulary may have a common core, they can vary according to register (Nevalainen & Tiekken-Boon van Ostade 2006: 303; Biber & Gray 2013). In order to see whether the use of phrasal verbs varies across registers, the tokens obtained from the corpus have been plotted in a boxplot (figure 4), which displays graphically the location and spread of a variable and provides some indication of data symmetry and skewness (see further Brezina 2018: 22–4).¹¹

The boxplot shows that phrasal verbs are unequally distributed across the various registers represented in CED. The value of their medians (marked by the horizontal black line) is remarkably different, with fiction presenting the highest median. The application of the Kruskal–Wallis test, however, leads to non-significant results ($p = 0.473$), probably because all the registers represent dialogic exchanges where phrasal verbs are common (unlike in formal, written documents). The two circles above the whiskers in dramatic texts and witness depositions represent two outliers (i.e. extreme values which are far from other values), which we must bear in mind during data analysis.

Figure 5 plots the overall normalized frequencies across registers in a bar line graph.¹² The highest frequencies of phrasal verbs are found in trial proceedings, followed by fictional texts and witness depositions. Conversely, the lowest rates occur in the two

¹¹ The boxplot was made based on the normalized frequencies of tokens of phrasal verbs per individual file.

¹² Raw figures are as follows: Didactic works (language teaching): 330; Didactic works (other): 733; Drama: 1,271; Fiction: 1,610; Trial proceedings: 2,243; Witness depositions: 1,102.

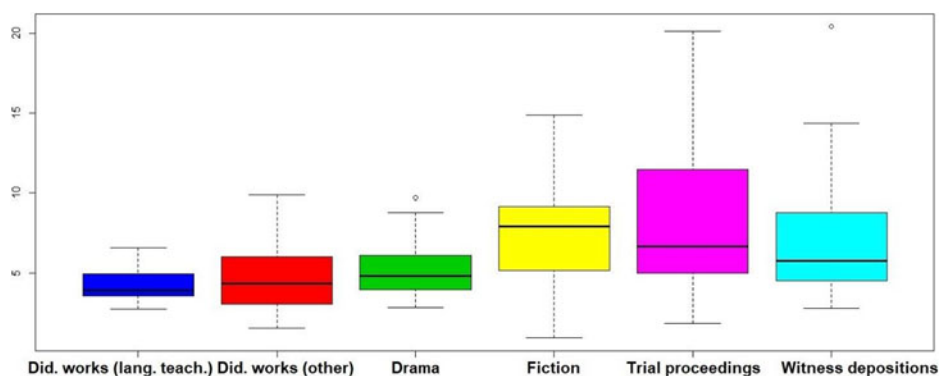


Figure 4. Distribution of phrasal verbs across registers in CED. Boxplot analysis

groups of didactic works and dramatic texts. In general, these figures indicate that phrasal verbs are more common in those categories which contain authentic, rather than constructed, dialogue, fiction being the only exception to this general tendency. We acknowledge, however, that this is a perception not supported by the statistical analysis.

Our results can be compared with those of Hiltunen (1994) and Claridge (2000), whose studies contain written, formal documents. Hiltunen (1994) analyzes a narrower set of particles (*away, back, forth, down, off, out* and *up*) and presents his results in percentages (1994: 136). For a more accurate comparison with our data, we normalized Hiltunen's raw figures in table 3, which also includes Claridge's (2000) data from the *Lampeter Corpus* and Thim's (2006b) results for the personal letters of CEEC.¹³ We cannot draw comparisons with Blake (2002) for several reasons. First, his study is not quantitative, and thus there are no frequencies to compare with ours here. Second, there are certain methodological differences. Blake's study is particularly criticized by Thim (2006a) for its inability to account for the occurrence of phrasal verbs in passages of formal poetry (Blake 2002: 37). However, Thim (2006a) fails to see that the combinations that Blake analyzes under the label 'phrasal verb' also include prepositional verbs and other structures containing prepositional phrases (e.g. *scorn'd at me* and *feared of all*).

As can be seen, in Hiltunen's data from the HC, the lowest frequency of phrasal verbs occurs in the most formal and authoritative types of documents, namely official letters and statutes, the successors of fifteenth-century Chancery English and representatives of the evolving standard norm at the beginning of the sixteenth century (Nevalainen & Raumolin-Brunberg 2011). The frequencies of phrasal verbs in those two text types are much lower than those found for all the registers analyzed in CED, which lends support to the idea that phrasal verbs were already colloquial and related to spoken

¹³ Thim's (2006a) data are not included in this table because he does not provide total raw figures, and frequencies of phrasal verbs are rendered in percentages rather than normalized. The normalized frequency for the personal letters of the CEEC (Thim 2006b) has been calculated by us on the basis of the data provided in his work.

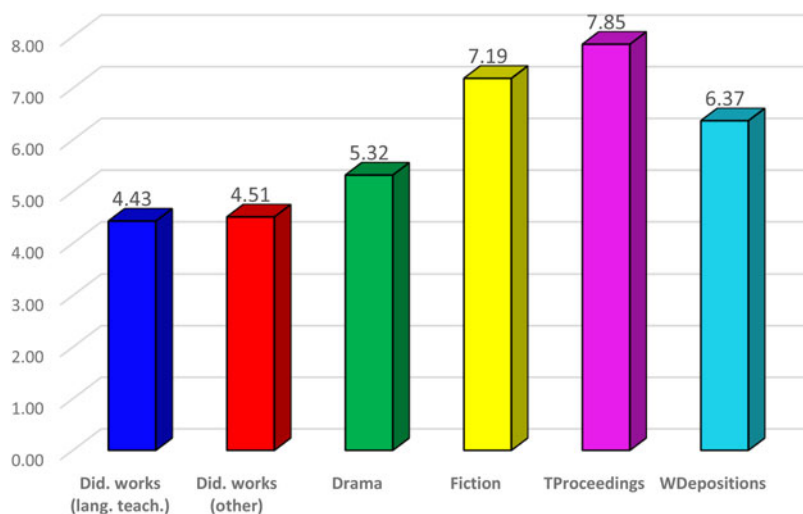


Figure 5. Distribution of tokens of phrasal verbs across registers in CED (normalized frequencies per 1,000 words)

Table 3. *Raw and normalized frequencies (per 1,000 words) of phrasal verbs in the EModE period (data from Hiltunen 1994, Claridge 2000 and Thim 2006b)*

Hiltunen (1994)	NF	Claridge (2000)	NF	Thim (2006b)	NF
Bible	3.36	Economy	4.6	Personal letters	2.16
Fiction	4.24	Law	3.8		
Handbooks	5.43	Miscellaneous	4.3		
Official letters	1.40	Politics	2.9		
Personal letters	2.77	Religion	3.2		
Sermons	2.57	Science	3.4		
Statutes	1.65				
Trial proceedings	2.32				

registers in the EModE period. Claridge's (2000) results for the category 'Law' are, however, much higher, but this is mostly due to the fact that this category in the *Lampeter Corpus* contains not only law documents, but also other types of texts thematically related to legal matters, such as trial proceedings, letters, treatises and even a document portraying a fictional dialogue (Claridge 2000: 190). In fact, all the categories in the *Lampeter Corpus* contain a mixture of different types of documents. This makes the interpretation of the results difficult, which is probably the reason why Thim (2006a: 299) finds Claridge's figures inconclusive and interpreted from the preconceived position that phrasal verbs are colloquial.

In contrast, the highest figures of phrasal verbs in Hiltunen's data are recorded in handbooks, fictional and biblical texts. As regards handbooks, it must be noted that the

handbooks in the HC are of a mixed nature, containing both dialogic and non-dialogic texts, which again makes their analysis problematic. As to fictional and biblical texts, the high frequency of phrasal verbs, though still lower than that of the fictional texts in CED, can be accounted for by the fact that both types of texts combine narratorial sections in the past with conversational exchanges between characters. As observed by Hiltunen (1994: 138) and Rodríguez-Puente (2019: 228 *et passim*), phrasal verbs, particularly those with literal meanings, are a very useful device to describe the events (both backgrounded and foregrounded eventive clauses)¹⁴ in the narrative sections. This specific function of phrasal verbs, illustrated in (24), is what may account for the fact that fictional texts present the second highest frequency of these combinations in Hiltunen's data and in CED, even higher than in dramatic texts. However, many combinations are also found in the conversational exchanges between the various characters, as shown in (25) and (26).

- (24) [...] but **flinging off** the Straw Hat, and **taking off** the Night-Rail, **away runs** he out at the Back-door, as if Nick had been at his Tail. (*Whole Pleasures of Matrimony*, 1688)
- (25) I vtterly deny: and doe (at this houre of my death) desire that heauen would **powre down** vengeance vpon me. (*Westward for Smelts*, 1620)
- (26) Do you not yet **bear away** with you some of the things that then you were conversant withal? (*Pilgrim's Progress*, 1678)

In contrast to biblical texts, the sermons of the HC have a rather low frequency of phrasal verbs, despite being speech-like texts. This is probably because sermons, although designed to be read aloud, are 'based on a careful written draft (and thus potentially influenced by the norms of written language)' (Claridge & Wilson 2002: 33), in which phrasal verbs are not, in principle, expected. Claridge (2000 188–90), on a closer inspection of the category 'Religion', observes that the sermons in the *Lampeter Corpus* contain a higher frequency of phrasal verbs than other religious texts. The frequency of phrasal verbs in the sermons of the *Lampeter Corpus* is slightly higher than that of the HC, possibly because of the time period covered by the two corpora. As phrasal verbs increase in frequency from EModE onwards, it is not surprising to find higher figures in the *Lampeter Corpus*, which covers the years 1640–1740.

The results for personal letters in Hiltunen's data may seem, a priori, 'unexpectedly low,' as Claridge describes them (2000: 187). They contain a very low frequency of phrasal verbs, a figure which is close to that found by Thim (2006b) in CEEC (see table 3). However, if we think of personal letters in earlier stages of the language, the low rates of phrasal combinations may turn out to be less surprising. Though portraying speech-like exchanges, EModE personal letters were not colloquial at all. They were carefully crafted through a conscious process of writing and followed specific conventions learned at home, in grammar schools and in letter-writing manuals (see, e.g. Anderson & Ehrenpreis 1966: 273; Austin 1998: 323; Nevalainen 2001). In fact, personal letters have changed considerably over

¹⁴ For the distinction between 'background' and 'foreground', see, among others, Hopper (1979).

time; whereas eighteenth-century letters were found to be expository, descriptive or argumentative in purpose, PDE letters are personally involved and interactive (see Biber & Finegan 1989, 1997; Biber 2001: 105). Indeed, as noted by Rodríguez-Puente (2019: 235–6), the frequency of phrasal verbs in personal letters increases significantly over time, particularly from the mid-nineteenth century onwards.

As shown in figure 5 above, the highest frequency of phrasal verbs in CED occurs in trial proceedings, which comes as no surprise, considering that they are based on real (not invented) speech, and hence linguistically close to spoken face-to-face interaction, not pre-planned and highly interactive (Culpeper & Kytö 2010: 63–4). Although the exchanges are produced in a formal setting, the proceedings contain passages which are quite colloquial in tone, some even recording bad language and insults (see Widlitzki & Huber 2016). Moreover, unlike witness depositions, whose narrative is normally in indirect speech, trial proceedings typically include transcripts in dialogue format (Culpeper & Kytö 2010: 49–59). By contrast, in Hiltunen's study trial proceedings display a much lower rate of phrasal verbs. This may be due to the difference in the number of particles analyzed or, more probably, the fact that the category 'Trial proceedings' of the HC also includes witness depositions, so that it is not entirely comparable with the samples of CED. Claridge (2000: 190–2) makes a more in-depth analysis of the real dialogue represented in three texts containing trials: one narrated in direct speech, another in indirect speech and a third which mixes dialogues and monologues (2000: 190). She points out that the highest figures of phrasal verbs in the *Lampeter Corpus* are found in a text containing transcriptions of direct speech produced by people from low socio-economic sectors which, for her, 'speaks for the fact that these verbs are indeed typical of spoken language, and thus probably also in general of more colloquial styles in the language' (2000: 192). Similar results are reported by Rodríguez-Puente (2019: 264–70) in comparing the trial proceedings of the *Old Bailey Corpus* with other registers represented in *A Corpus of English Historical Registers* (ARCHER). Therefore, a closer inspection of the results in previous studies and the data obtained from the larger sample of CED here seems to confirm that EModE phrasal verbs are indeed a feature of the spoken language, as they are amply represented in speech-like texts, especially those which contain authentic dialogue with minimal narratorial intervention.

Some illustrative examples of how phrasal verbs are used in the spoken language of trial proceedings are presented in (27) to (30). Most of these examples contain features of subjectivity and personal affect, typical of oral interaction, such as the use of first and second person pronouns, private verbs (e.g. *think* and *conceive*), direct questions (see Biber 1988: 225; Taavitsainen 1994: 202) and, as shown, also phrasal verbs.

- (27) Can Mankind be persuaded, that you, that had this Negotiation in 74. and 75. **left off** just then [...]. (*Tryal of Edward Coleman*, 1678)
- (28) I think it is so now: Now you are come to be arraigned, and the evidence is to **come in** against you [...]. (*Triall of Connor Lord Macguire*, 1644)

- (29) And then when I went into the House of Lords, I **made out** my charge against him. (*Tryals of Robert Green*, 1678/9)
- (30) Frances Countess of Somerset, **hold up** thy Hand. (*Trial of the Lady Frances*, 1616)

Interestingly, although produced in a formal context, trial proceedings acquire a rather colloquial tone in certain passages, especially during the testimonies of witnesses and the accused, where insults and bad language are occasionally found alongside phrasal verbs (see (31)–(33)).

- (31) [...] Annesley **came up** with his Gun, and swore, God damn your Blood, deliver your Net, or you are a dead Man; and he **fir'd off** before he received any Answer from my father. (*Trial of James Annesley and Joseph Redding*, 1742)
- (32) [...] you have good reason to disclaim them with their bloody religion, and to **seek out** better advisers for you at your death [...]. (*Triall of Connor Lord Macguire*, 1644)
- (33) Because then I shall heare popish doctrine; I was once in the whore's bosome, and these hornes **thrust me in**. (*High Commission*, 1632)

In contrast to trial proceedings, witness depositions are normally narrated in indirect speech, as in (34), where phrasal verbs are integrated in the narration, being particularly useful for the description of events.

- (34) [...] Joan Buts **came in**, and **sat her down** upon a Stool, looking with a most frightful and gastly Countenance; and being asked by a Woman that was there present, what she ailed? she answered, She was not well, nor had been out in seven weeks before; why would you **come out** now then, said the Woman I could not forbear coming to see you, said she; and with that, she **threw down** her Hat and **tumbled down**, wallowing on the ground, making a fearful and dismal noise; and being **got up**, she fell a cursing in a most horrid manner. (*Examination of Joan Buts*, 1682)

Narrational sections are not so frequent in didactic works and drama. Dramatic texts include some asides, but these tend to be extremely short and contain few phrasal verbs, most of them indicating movement (e.g. *come in* and *go out*). The scarcity or complete lack of such narratorial sections is perhaps what explains the lower rates of phrasal verbs in these registers (especially in contrast to fiction), which otherwise contain combinations that provide their conversational exchanges with a colloquial tone, as in (35) to (37). It must be noted, however, that these texts represent constructed dialogue and, even though authors may decide to represent all kinds of characters, their way of speaking and interacting is crafted rather than natural, which may account for the relatively lower use of phrasal verbs in them.

- (35) Yea I will quarter him, and pull all the bones out of his flesh, then will I **barrell vp** his bowels. (*Menaecmi*, 1595)
- (36) the Lawyers they **pocket vp** all the gaines. (*Menaecmi*, 1595)
- (37) **Pluck it up** woman do not fear. (*Honest John and Loving Kate*, 1685)

We move on now to the analysis of the diachronic distribution of phrasal verbs across the various registers, with the prediction that their evolution may not have been the same. As is well-known, the standardization of English involved a change within the written language which, particularly during the eighteenth century, rendered it less oral and more literate, so that it would be more suitable for the manifold functions of the standard language. Extensive research by Biber (2001) and Biber & Finegan (1989, 1992, 1997) on the development of registers from the seventeenth to the twentieth century in ARCHER has consistently demonstrated that styles tend to become more literate during the eighteenth century, and then registers drift apart linguistically, some becoming more oral and others more literate. McIntosh (1998: 23–4) accounts for the change towards a more polished, written-like style in the eighteenth century in terms of standardization, gentrification, the cleaning-up and modernization of English, and the culmination of the prescriptivist period.

It seems necessary, then, to see whether differences can be detected in this respect when we compare the samples of authentic dialogues with those containing constructed dialogues. In the former, such a turn towards more literate styles is, in principle, not expected. The latter, however, represent dialogues crafted by authors, who might be expected to be more predisposed to adopt the stylistic fashion of the time. The evolution of phrasal verbs across registers is represented in figure 6, which clearly confirms our initial predictions.

Towards the end of the sixteenth century all the registers represented in CED have frequencies higher than 3.5, that is, much higher than the mean of the most formal, written registers of the HC (i.e. statutes and official letters; see table 3), which implies that these combinations were clearly a feature of the spoken language of the time. During the first corpus subperiod, fiction is the register with the highest frequency of phrasal verbs, a rate which remains quite high until subperiod 1600–39, but then begins to decrease progressively until the last corpus subperiod ($p = 0.1717$). A similar, though not as marked, path is followed by other registers containing constructed dialogues, namely drama ($p = 1$) and didactic works ‘Other’ ($p = 0.9143$), where frequencies remain rather stable from the first to the last subperiod. In turn, didactic works on language teaching have higher rates of phrasal verbs in the last subperiod than in the first, though from subperiod 1680–1719 their frequency evolves slightly downwards as well ($p = 0.6667$). At the opposite end of the spectrum, trial proceedings contain the lowest rates of phrasal combinations at the beginning of the period, whereas witness depositions present frequencies similar to those of dramatic texts. However, in trial proceedings the frequency of phrasal verbs grows significantly from the first to the last subperiod ($p = 0.01714$), reaching its peak in subperiod 1720–60. Witness depositions still present relatively low rates of phrasal verbs in the second subperiod, but these grow steadily from then on ($p = 0.0303$). The sharp and significant escalation in the use of phrasal verbs in registers that contain real language contrasts strongly with the stability seen in registers that contain constructed dialogue. If phrasal verbs were growing in number, they should be growing across all registers equally. Our hypothesis here is

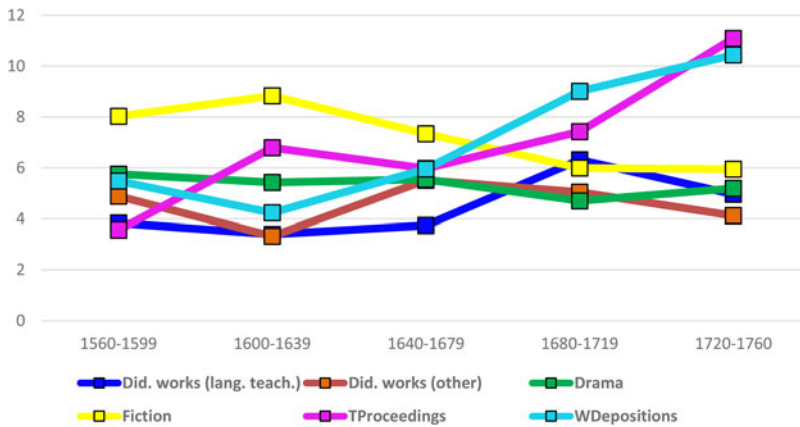


Figure 6. Diachronic distribution of phrasal verbs across registers CED (normalized frequencies per 1,000 words)

that the change towards more literate styles in the eighteenth century may have affected those registers containing invented speech. These do not present increasing figures of phrasal verbs, perhaps as a reflection of their authors seeking to maintain a more polished and formal style.

A striking finding from the data, as presented in figure 6, is the extremely low frequency of phrasal verbs in the trial proceedings of the first subperiod. In order to try to account for this, we looked at the dispersion of our data in this specific register, shown in figure 7, where the y -axis represents the normalized frequencies of individual texts and the x -axis the corpus subperiods. Although no outliers were present in the group of trial proceedings in the boxplot analysis (figure 4), the dispersion graph shows that the three trials in the first subperiod contain very low frequencies of phrasal verbs, whereas in the second subperiod all the trials have frequencies higher than five. In the remaining periods there is considerable variety, with some texts situated below five and the highest being above twenty.

In order to ascertain why the frequencies of phrasal combinations vary so drastically within the same period, we read the individual texts to see whether their subject matter might be a reason, as some authors have argued (Thim 2006a: 302; Rodríguez-Puente 2019: 224). Rather than the subject matter, however, we suspect that this variation is due to the specific style used by the transcribers of the trials or the participants involved in them. Trial proceedings are supposed to portray conversational exchanges with minimal narratorial intervention, but we cannot fully discard the possibility of scribal manipulation or stylistic bias. The proceeding with the lowest frequency in the corpus is the *Trial of Mr. Robert Hickford* (1571). Although this text contains conversational exchanges in first-person narration, these are very formal in tone, either because the speech itself was pre-planned by the speaker or because it was rendered as such by the scribe. Note, for example, the highly nominalized discourse illustrated in (38), more typical of formal written language than of natural conversation.

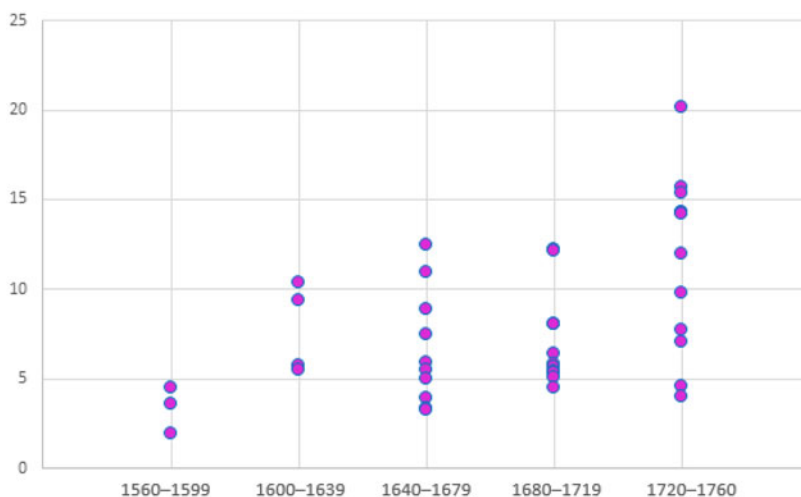


Figure 7. Dispersion of phrasal verbs in the trial proceedings of CED (normalized frequencies per 1,000 words)

(38) I know the Law hath not Intent to the Conscience, or Intent of Men; but is to judge only of the Mind, according to **the appearing of outward Facts**: so the Law cannot accept my Intent for **my Purgation**. Therefore I shall declare **the Truth of my Doing**, and upon that further open unto you at large **my Dealing** in the Matters contain'd in the Indictment. First, For **the Matter of moving of Sedition**, I did never know that any such thing was meant by my Lord [...].

The formality of this statement contrasts strongly with the short question and answer exchanges which characterize the trial in which we find the highest frequency of phrasal verbs, namely *Minutes taken at the Court Martial, held upon Captain John Ambrose* (1745). Extract (39) illustrates that phrasal verbs here are mostly literal, used to describe the movements of the ships, and are also quite repetitive, as they appear first in the questions and are then sometimes repeated in the answers.

(39) *Q.* Supposing the Rupert had **bore down** in a Line with the Admiral, and **brought up** in a Line when he did, how long would it have been before the Rupert would have had any Ship to engage?
A. If the Rupert had **gone down** and **brought up** in a Line with the Admiral, she would have been in a proper Position to engage the five Sail of the Enemy **coming up**.
Q. Did the Rupert stop those Ships from **getting up**?

Although repetitions are part of natural conversation, in this particular case the subject matter of the text (the description of the movements of ships) is also relevant in accounting for the high frequency of phrasal verbs.

5 Conclusion

Based on a sample of over 7,000 corpus examples, in this article we have demonstrated that phrasal verbs were well entrenched in the spoken, colloquial language of the EModE period. The wide variety of combinations, the attestation of non-compositional meanings and the ample number of *hapax legomena* all serve to underline the fact that phrasal verbs were solidly consolidated at the time. Moreover, the analysis of the diachronic distribution of phrasal combinations shows that the growth of phrasal verbs in CED begins at a rather early stage (already from 1600 onwards), much earlier than usually reported in the literature (1700 onwards). The growth of phrasal verbs is particularly marked in those registers which most closely represent the spoken language of the past, namely trial proceedings and witness depositions, especially towards the eighteenth century. Arguably, this happens because, contrary to those texts which represent constructed dialogues, they are not conditioned by the stylistic fashion of the time which promoted the adoption of a more literate style.

Crucially, this quantitative diachronic study has allowed us to describe empirically and objectively phrasal verbs in the EModE period. In so doing, we revisited earlier studies which were seen to be ‘meaningless’ by Thim (2012: 213), noting that the proportion of phrasal verbs in the texts of CED is much larger than that found in the most authoritative and formal written types of documents, namely statutes and official letters. Moreover, we provide an alternative explanation for the apparently chaotic distribution of phrasal verbs which originally led Thim to propose his ‘colloquialization conspiracy’, and we point to methodological differences, as well as differences among registers which were probably not considered in earlier work (e.g. the non-colloquial character of EModE personal letters). Once more, we demonstrate the importance of register analysis for the study of the development of languages while acknowledging that other features, such as subject matter, as well as the idiosyncrasies and personal style of authors and/or speakers, can also trigger or deflect the use of phrasal verbs. These new insights into the history of phrasal verbs through the data presented here lead us to conclude that the colloquial status of phrasal verbs at the time cannot be understood merely in terms of a ‘colloquialization conspiracy’.

Authors' addresses:

Dept of English, French and German Philology

University of Oviedo

Campus El Milán

C/ Amparo Pedregal s/n. 33011

Oviedo

Asturias

Spain

rodriguezppaula@uniovi.es

mariaobaya@gmail.com

References

- Akimoto, Minoji. 2006. On the decline of *after* and *forth* in verb phrases. In Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds.), *Syntax, style and grammatical norms: English from 1500–2000*, 11–31. Bern: Peter Lang.
- Anderson, Howard & Irvin Ehrenpreis. 1966. The familiar letter in the eighteenth century: Some generalizations. In Howard Anderson, Philip B. Daghlian & Irvin Ehrenpreis (eds.), *The familiar letter in the eighteenth century*, 269–82. Lawrence, KS: University of Kansas Press.
- Austin, Francis. 1998. Epistolary conventions in *The Clift Family Correspondence*. In Mats Rydén, Ingrid Tieken-Boon van Ostade & Merja Kytö (eds.), *A reader in Early Modern English*, 319–47. Frankfurt am Main: Peter Lang.
- Baayen, Harald. 1989. A corpus-based approach to morphological productivity: Statistical analysis and psycholinguistic interpretation. PhD dissertation, Free University of Amsterdam.
- Baayen, Harald. 1992. Quantitative aspects of morphological productivity. In Geerd Booij & Jappa van Marle (eds.), *Yearbook of morphology 1991*, 109–49. Dordrecht: Kluwer Academic Publishers.
- Baayen, Harald. 1993. *Word frequency distributions*. Dordrecht: Kluwer Academic Publishers.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, Douglas. 2001. Dimensions of variation among eighteenth-century speech-based and written registers. In Diller & Görlach (eds.), 89–109.
- Biber, Douglas. 2003. Variation among university spoken and written genres: A new multi-dimensional analysis. In Pepi Lestyna & Charles F. Meyer (eds.), *Corpus analysis. Language structure and language use*, 47–70. Amsterdam: Rodopi.
- Biber, Douglas & Edward Finegan, 1989. Drift and the evolution of English style: A history of three genres. *Language* 65(3), 487–517.
- Biber, Douglas & Edward Finegan. 1992. The linguistic evolution of five written and speech-based genres from the seventeenth to the twentieth centuries. In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen & Irma Taavitsainen (eds.), *History of Englishes: New methods and interpretations in historical linguistics*, 688–704. Berlin: Mouton de Gruyter.
- Biber, Douglas & Edward Finegan. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 253–75. Helsinki: Société Néophilologique.
- Biber, Douglas & Bethany Gray. 2013. Being specific about historical change: The influence of sub-register. *Journal of English Linguistics* 41(2), 104–34.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Blake, Norman Francis. 2002. Phrasal verbs and associated forms. *Atlantis* 24(2), 25–39.
- Bolinger, Dwight. 1971. *The phrasal verb in English*. Cambridge, MA: Harvard University Press.
- Brezina, Vaclav. 2018. *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press.
- Brinton, Laurel J. 1988. *The development of English aspectual systems: Aspectualizers and post-verbal particles*. Cambridge: Cambridge University Press.
- Brown, David West & Chris C. Palmer. 2015. The phrasal verb in American English: Using corpora to track down historical trends in particle distribution, register variation, and noun collocations. In Michael Adams, Laurel J. Brinton & Roger D. Fulk (eds.), *Studies in the history of the English language VI: Evidence and method in histories of English*, 71–97. Berlin: Mouton de Gruyter.
- Castillo, Concha. 1994. Verb-particle combinations in Shakespearean English. *Neuphilologische Mitteilungen* 95(4), 439–51.

- Claridge, Claudia. 2000. *Multi-word verbs in Early Modern English: A corpus-based study*. Amsterdam: Rodopi.
- Claridge, Claudia & Andrew Wilson. 2002. Style evolution in the English sermon. In Teresa Fanego, Belén Méndez-Naya & Elena Seoane (eds.), *Sounds, words, texts and change. Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000*, 25–44. Amsterdam: John Benjamins.
- Cowie, Anthony P. & Ronald Mackin. 1975. *Oxford dictionary of current idiomatic English*, vol. I: *Verbs with prepositions and particles*. Oxford: Oxford University Press.
- Culpeper, Jonathan & Merja Kytö. 2010. *Early Modern English dialogues. Spoken interaction as writing*. Cambridge: Cambridge University Press.
- Cusack, Brigit (ed.). 1998. *Everyday English 1500–1700: A reader*. Edinburgh: Edinburgh University Press.
- De Smet, Hendrik. 2010. Grammatical interference: Subject marker ‘for’ and the phrasal verb particles ‘out’ and ‘forth’. In Elizabeth C. Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 75–104. Amsterdam: John Benjamins.
- Dempsey, Kyle B., Philip M. McCarthy & Danielle S. McNamara. 2007. Using phrasal verbs as an index to distinguish text genres. *Natural Language Processing* 13(3), 217–22.
- Denison, David. 1981. Aspects of the history of English group-verbs, with particular attention to the syntax of the Ormulum. PhD dissertation, Oxford University.
- Denison, David. 1985. The origins of completive *up* in English. *Neuphilologische Mitteilungen* 86 (1), 37–61.
- Denison, David. 1998. Syntax. In Suzanne Romaine (ed.), *The Cambridge history of the English language*, vol. IV: 1776–1997, 92–329. Cambridge: Cambridge University Press.
- Diller, Hans-Jürgen & Manfred Görlach (eds.). 2001. *Towards a history of English as a history of genres*. Heidelberg: Universitätsverlag C. Winter.
- Fraser, Bruce. 1976. *The verb-particle combination in English*. New York: Academic Press.
- Hiltunen, Risto. 1994. Phrasal verbs in Early Modern English: Notes on lexis and style. In Dieter Kastovsky (ed.), *Studies in Early Modern English*, 129–40. Berlin: Mouton de Gruyter.
- Hiltunen, Risto. 1999. Verbal phrases and phrasal verbs in Early Modern English. In Laurel J. Brinton & Minoji Akimoto (eds.), *Collocational and idiomatic aspects of composite predicates in the history of English*, 133–65. Amsterdam: John Benjamins.
- Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In Talmy Givón (ed.), *Discourse and syntax* (Syntax and Semantics 12), 213–41. New York: Academic Press.
- Huber, Magnus. 2007. The *Old Bailey Proceedings*, 1764–1834. Evaluating and annotating a corpus of 18th- and 19th-century spoken English. In Anneli Meurman-Solin & Arja Nurmi (eds.), *Annotating variation and change* (Studies in Variation, Contacts and Change in English 1). Helsinki: University of Helsinki. <https://varieng.helsinki.fi/series/volumes/01/huber/>
- Ishizaki, Yasuaki. 2009. A usage-based analysis of the distribution of *forth* in the history of English. *Kindai Eigo Kenkyu* (Studies in Modern English 25), 41–61.
- Kennedy, Arthur Garfield. 1920. *The modern English verb-adverb combination*. Stanford, CA: Stanford University Press.
- Konishi, Tomoshichi. 1958. The growth of the verb-adverb combination in English – A brief sketch. In Kazuo Araki (ed.), *Studies in English grammar and linguistics: A miscellany in honour of Takanobu Otsuka*, 117–28. Tokyo: Kenkyusha.
- Kytö, Merja & Jonathan Culpeper. 2006. *A Corpus of English Dialogues 1560–1760*. Compiled under the supervision of Merja Kytö (Uppsala University) and Jonathan Culpeper (Lancaster University).

- Kytö, Merja & Erik Smitterberg. 2006. Nineteenth-century English: An age of stability or a period of change? In Roberta Facchinetti & Matti Rissanen (eds.), *Corpus-based studies of diachronic English*, 199–230. Bern: Peter Lang.
- Kytö, Merja & Terry Walker. 2006. *Guide to A Corpus of English Dialogues 1560–1760*. Uppsala: Acta Universitatis Upsaliensis.
- Liu, Dilin. 2011. The most frequently used English phrasal verbs in American English: A corpus-based analysis. *TESOL Quarterly* 45(4), 661–88.
- Los, Bettelou. 2004. From resultative predicate to event-modifier. The case of *forth* and *on*. In Christian Kay, Simon Horobin & Jeremy Smith (eds.), *New perspectives on English historical linguistics*, vol. I: *Syntax and morphology*, 83–102. Amsterdam: John Benjamins.
- Makkai, Adam. 1972. *Idiom structure in English*. The Hague: Mouton.
- Martin, Pamela. 1990. The phrasal verb: Diachronic development in British and American English. PhD dissertation, Columbia University Teachers College.
- McArthur, Tom. 1989. The long-neglected phrasal verb. *English Today* 18, 38–44.
- McIntosh, Carey. 1998. *The evolution of English prose 1700–1900: Style, politeness, and print culture*. Cambridge: Cambridge University Press.
- Nevalainen, Terttu. 1999. Early Modern English lexis and semantics. In Roger Lass (ed.), *The Cambridge history of the English language*, vol. III: 1476–1776, 332–458. Cambridge: Cambridge University Press.
- Nevalainen, Terttu. 2001. Continental conventions in Early English correspondence. In Diller & Görlach (eds.), 203–25. Heidelberg: Universitätsverlag C. Winter.
- Nevalainen, Terttu & Ingrid Tieken-Boon van Ostade. 2006. Standardisation. In Richard Hogg & David Denison (eds.), *A history of the English language*, 271–311. Cambridge: Cambridge University Press.
- Nevalainen, Terttu & Helena Raumolin-Brunberg. 2011. *Helsinki Corpus*. Period division: Early Modern English. Retrieved 16 December, 2021 from www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus/earlymodern2.html
- Oxford English Dictionary*. Third edition in progress: *OED Online*, March 2000–. Ed. Michael Proffitt. www.oed.com
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- R Core Team. 2020. *R version 4.0.3. A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. www.R-project.org
- Rodríguez-Puente, Paula. 2012. The development of non-compositional meanings in phrasal verbs: A corpus-based study. *English Studies* 93(1), 71–90.
- Rodríguez-Puente, Paula. 2013. A turn of the screw in the semantics of phrasal verbs: Phrasal verbs with *up* as a test case. In Irén Hegedüs & Dóra Pödör (eds.), *Periphrasis, replacement and renewal: Studies in English historical linguistics*, 243–65. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Rodríguez-Puente, Paula. 2017. Tracking down phrasal verbs in the spoken language of the past: Late Modern English in focus. *English Language and Linguistics* 21(1): 69–97.
- Rodríguez-Puente, Paula. 2019. *The English phrasal verb, 1650–present: History, stylistic drifts, and lexicalisation*. Cambridge: Cambridge University Press.
- Scott, Mike. 2020. *WordSmith Tools*, version 8. Stroud: Lexical Analysis Software.
- Smitterberg, Erik. 2008. The progressive and phrasal verbs: Evidence of colloquialization in nineteenth-century English? In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta & Minna Korhonen (eds.), *The dynamics of linguistic variation. Corpus evidence on English past and present*, 269–89. Amsterdam: John Benjamins.
- Spasov, Dimiter. 1966. *English phrasal verbs*. Sofia: Nauka Izkoustvo.

- Taavitsainen, Irma. 1994. Subjectivity as a text-type marker in historical stylistics. *Language and Literature* 3(3), 197–212.
- Thim, Stefan. 2006a. Phrasal verbs in everyday English, 1500–1700. In Andrew James Johnston, Ferdinand von Mengden & Stefan Thim (eds.), *Language and text: Current perspectives on English and Germanic historical linguistics and philology*, 291–306. Heidelberg: Winter.
- Thim, Stefan. 2006b. Phrasal verbs in Late Middle and Early Modern English: Combinations with *back*, *down*, *forth*, *out* and *up*. In Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds.), *Syntax, style and grammatical norms: English from 1500–2000*, 213–28. Bern: Peter Lang.
- Thim, Stefan. 2011. On the phrasal verbs in some Paston Letters. In Renate Bauer & Ulrike Krischke (eds.), *More than words: English lexicography and lexicology past and present. Essays presented to Hans Sauer on the occasion of his 65th birthday*, part I, 355–85. Bern: Peter Lang.
- Thim, Stefan. 2012. *Phrasal verbs: The English verb-particle construction and its history*. Berlin: Mouton de Gruyter.
- Widlitzki, Bianca & Magnus Huber. 2016. Taboo language and swearing in 18th century and 19th century English: A diachronic study based on the *Old Bailey Corpus*. In María José López-Couso, Belén Méndez-Naya, Paloma Núñez-Pertejo & Ignacio M. Palacios-Martínez (eds.), *Corpus linguistics on the move: Exploring and understanding English through corpora*. Amsterdam: John Benjamins, 313–36.
- Wild, Kate. 2010. Attitudes towards English usage in the Late Modern Period: The case of phrasal verbs. PhD dissertation, University of Glasgow.
- Yáñez-Bouza, Nuria. 2015. *Grammar, rhetoric and usage in English: Preposition placement 1500–1900*. Cambridge: Cambridge University Press.