

## Parenthetical reporting clauses in the history of English: the development of quotative inversion

ANNA CICHOSZ

University of Łódź

(Received 29 August 2017; revised 4 December 2017)

This study is a corpus-based diachronic analysis of English reporting parentheticals, i.e. clauses introducing direct speech, placed after or in the middle of the reported message. The aim of the investigation is to trace the development of the construction throughout the history of English, establishing the main factors influencing the choice between VS and SV patterns (i.e. with and without quotative inversion respectively), showing how various reporting verbs were increasingly attracted to the construction, and demonstrating the gradual morphological reduction of the main reporting verbs: *quoth* and *say*. The study is based on syntactically annotated corpora of Old, Middle, Early Modern and Late Modern English, and uses other corpora to illustrate more recent changes. The study reveals that reporting clauses do not show regular quotative inversion with all subject types until the Early Modern English period and links this development to the emergence of the comment clause with *say*. It is also claimed that quotative inversion is not directly derived from the V-2 rule and that parenthetical reporting clauses have functioned as a separate construction since the Old English period.

**Keywords:** parenthetical clause, reporting clause, reporting verb, quotative inversion, English syntax

### 1 Reporting clauses and quotative inversion

In Present-day English (PDE), a quotation may be introduced into the discourse in various ways. Numerous studies of English reporting structures show the relatively recent development of innovative quotatives such as *be like* or *this is me/him* (e.g. Buchstaller 2013; Lampert 2013; Barbieri 2005; Tagliamonte & D’Arcy 2004), as in (1) and (2).

(1) **I was like**, ‘But I won’t’ (Tagliamonte & D’Arcy 2004: 493)

(2) **this is me** ‘what...what’s your ... what’s your problem?’ (Cheshire & Fox 2008)

Nonetheless, traditional quotatives in the form of reporting clauses (i.e. clauses with a verb of saying whose function is to introduce direct speech) are still the dominant strategy in written English and in more formal registers. Such a reporting clause may precede the quoted message, as in (3), it may follow it, as in (4), or be placed within the quotation, as in (5). The last two contexts (henceforth referred to as ‘parenthetical reporting clauses’ or ‘the parenthetical reporting construction’ abbreviated as ‘PRC’) are especially interesting from a syntactic point of view because they often display the so-called quotative inversion, i.e. subject–verb inversion in reporting clauses.<sup>1</sup>

<sup>1</sup> Very rarely a reporting clause with SV inversion may also precede the quoted message. The phenomenon is known as clause-initial quotative inversion; its use is restricted to ‘journalistic writing’ (Biber *et al.* 1999: 922;

- (3) **She said:** ‘Elderly people often have smaller groups of friends and family to support them’ (NEWS)
- (4) ‘That’s the whole trouble,’ **said Gwen**, laughing slightly. (FICT)
- (5) Sketching, **says Uderzo**, is a fast process. (NEWS)  
(Biber *et al.* 1999: 921–2)

Quotative inversion may only occur if the verb is in the simple present or simple past and it is ‘most common when the verb is *said*, the subject is not a pronoun, and the reporting clause is medial’ (Quirk *et al.* 1985: 1022). When the subject is a personal pronoun as in (6), the verb phrase is complex as in (7), or the clause shows expansion to the right with specification of the addressee as in (8), subject-verb order ‘is virtually the rule’ (Biber *et al.* 1999: 922).

- (6) ‘The safety record at Stansted is first class,’ **he said**. (NEWS)
- (7) ‘Konrad Schneider is the only one who matters,’ **Reinhold had answered**. (FICT)
- (8) There’s so much to living that I did not know before, **Jackie had told her happily**. (FICT)  
(Biber *et al.* 1999: 922)

Pronominal subjects show inversion very infrequently and the pattern illustrated in (9) is said to be archaic (Quirk *et al.* 1985: 1022; Huddleston & Pullum 2002: 1027).

- (9) ‘We may all be famous, then,’ **said he** (FICT) (Biber *et al.* 1999: 922)

Moreover, the fixed combination *says I* is used by some speakers to report a conversation, as in (10); the theoretically more proper form *say I* is never used in this context (Biber *et al.* 1999: 922). Quirk *et al.* (1985: 1024, fn. b) report that the non-standard structure shown in (10) is used in ‘old-fashioned speech’.<sup>2</sup>

- (10) I said I think something’s gone wrong with the auto bank machine, **says I** (CONV)  
(Biber *et al.* 1999: 922)

Thus, PDE quotative inversion is virtually restricted to nominal subjects, but even when such subjects are accompanied by simple verb forms with no specification of the addressee, they follow the non-inverted order quite regularly, as in (11).

- (11) ‘In those days,’ **Sue admitted**, ‘we were heavily in debt.’  
(Huddleston & Pullum 2002: 1027)

The variation between SV as in (11) and VS as in (4) and (5) is also influenced by genre: in fiction both orders are attested, though ‘with a slight preference for the regular SV order’, while in news inversion ‘is strongly preferred’ (Biber *et al.* 1999: 923).

Quirk *et al.* 1985: 1024, fn. c; Huddleston & Pullum 2002: 1027, fn. 35; Sams 2009: 168) and it is reported to be particularly frequent in American magazines, especially *Time* (Sonoda 1997; Konno 2013). This structure is not taken into account in the present study, which focuses on parenthetical (i.e. non-initial) reporting clauses.

<sup>2</sup> Quotative inversion is said to be restricted to the written register; Joseph (2000: 312) notes that ‘it has no direct correspondent in spoken usage’. However, the fact that examples such as (10) are attested in corpora of spoken English suggests that it is simply much less frequent in speech than in writing.

As far as the structural analysis of reporting clauses with quotative inversion is concerned, they are treated as ambiguous between V-initial and V-second because the reported message may (but does not have to) be interpreted as the object of the reporting clause (Biber *et al.* 1999: 921). As discussed by Quirk *et al.* (1985: 1022), ‘the structural relationship between the reporting clause and the direct speech poses some analytical problems’. In formal accounts, quotative inversion is separated from other types of inversion still attested in PDE though the analysis of the structure is a controversial issue (Collins & Branigan 1997 vs Bruening 2016). Nonetheless, it is suggested that English quotative inversion ‘may derive from the verb-second construction’ (Los 2009), which was attested in Old English (OE) and was ‘largely lost during the late Middle English period’ (Fischer *et al.* 2000: 105). Rohrbacher (1994: fn. 3) notes that since quotative inversion is restricted to the written register, it is possible ‘that it reflects an earlier stage of English and cannot be analyzed in purely synchronic terms’.

Scholars working on early English tend not to treat quotative inversion as a part of the V-2 rule. In various analyses of Middle English (ME) syntax focusing on the loss of the V-2 order, quotative inversion is mentioned but only to state that it would not be included in the investigation because of its distinct discourse characteristics (Los 2009; Westergaard 2007) and structural ambiguity, which make it different from typical V-2. Haeberli (2002) excludes such clauses from his investigation, giving the following example:

- (12) ‘Syre,’ **seide Moises**, ‘zif men aske how men clepeþ 3ow, what schal I seye?’  
 “‘Lord,” said Moses, “if somebody asks what you are called, what should I say?”  
 (Vices, 101.88)

According to Haeberli, the status of quotative inversion is not on a par with other V-2 clauses he analyses in his study because ‘[a]t the surface, [it] looks like a parenthetical V-1 clause rather than like a genuine example in which subject-verb inversion occurs due to the fronting of a non-operator’ (Haeberli 2002: 271). A similar view is expressed by Bech in the context of OE syntax: ‘It is therefore not obvious that the syntax of constructions containing direct speech has the integrated structure that enables syntactic-typological reasoning; it is closer to a text structure containing a sequence of two main clauses’ (Bech 2017: 12). Therefore, scholars working on OE usually regard clauses such as (13) as instances of the rare V-1 order (which is said to be particularly well attested in clauses with verbs of saying) (Calle-Martín & Miranda-García 2010; Mitchell 1985: §3930) and not as examples of V-2.

- (13) Min gemet is, **cwæp Paulus**, þæt ic bege mine cneowa.  
 My measure is said Paul that I bow my knees  
 ‘It is my measure, said Paul, that I bend my knees’ (coblick, LS\_32\_[PeterandPaul [BlHom\_15]]:187.285.2410)

On the other hand, some scholars see PDE quotative inversion as ‘apparently a residue of earlier English where V2 was much more pervasive’ (Zwart 2005: 19). In modern Germanic V-2 languages it is obligatory and associated with the V-2

phenomenon (Zwart 2005: 18–19; Griffiths 2015) or considered as a structure which is indistinguishable from other types of V-2 (Harbert 2007: 414), even though obligatory or optional quotative inversion is also attested in other, non-V-2 languages such as Hungarian (Gärtner & Gyuris 2014), Spanish (Suñer 2000; Matos 2013), Portuguese (Ambar 1992; Matos 2013), French (Bonami & Godard 2008) and Russian (Bailyn 2012: 318).

The aim of this study is to investigate the development of parenthetical reporting clauses throughout the history of English, with a special focus on the variation between SV and VS patterns and on the factors influencing it in particular stages of English. Section 2 describes the methodology of the study. In section 3, general quantitative results are presented for each period separately. Sections 4 and 5 focus on two main verbs used in the parenthetical reporting construction, *quoth* and *say*, showing their diachronic development. Section 6 is an account of the development of the investigated construction within the diachronic construction grammar framework, while section 7 presents the conclusions.

## 2 Methodology

The study was conducted on the basis of the *York–Toronto–Helsinki Parsed Corpus of Old English Prose* (YCOE, c. 1.5 million words; Taylor *et al.* 2003), the *Penn–Helsinki Parsed Corpus of Middle English* (PPCME2, c. 1.2 million words; Kroch & Taylor 2000), the *Penn–Helsinki Parsed Corpus of Early Modern English* (PPCEModE, c. 1.7 million words; Kroch *et al.* 2004) and the *Penn Parsed Corpus of Modern British English* (PPCMBE2, c. 2.8 million words; Kroch *et al.* 2016) searched by means of the *CorpusSearch 2* application (Randall *et al.* 2005–13).<sup>3</sup> In order to identify reporting clauses, first all parenthetical clauses with a lexical verb were extracted from the corpora (the query: \*PRN\* iDoms VB\*; all reporting clauses, both clause-medial and clause-final, are annotated in the corpora as parenthetical). Since the corpora are not lemmatised, it was necessary to go through the results manually and find parenthetical clauses containing reporting verbs. All clauses with a verb of saying and an overt subject were initially included in the sample, both with and without inversion, and also with some additional clause elements, mostly objects and adverbials. Punctuation marks were not taken into consideration in the analysis since their use in various editions of early English texts is inconsistent and cannot be treated as a reliable guide for identifying instances of direct speech; see (13) where no quotation marks are used. At this stage, some of the examples extracted from the corpora could be interpreted as comment clauses containing verbs of saying, i.e. clauses whose function

<sup>3</sup> In addition, for selected reporting verbs more recent data were extracted from the *Corpus of Contemporary American English* (COCA; Davis 2008), the *Corpus of Historical American English* (COHA; Davis 2010) and the spoken component of the *British National Corpus* (BNC; Coleman *et al.* 2012), as explained in the relevant sections of the article. The BNC was searched by means of *Spokes for the BNC* (Pezik 2015).

Table 1. *The composition of the final study sample*

	PRN with lexical verbs	PRN with verbs of saying and S <sup>4</sup>	1st-person present verbs	Final sample
Old English	1,086	209	13	196
Middle English	1,974	1,381	14	1,367
Early Modern English	4,464	1,405	114	1,291
Late Modern English	4,332	1,156	138	1,018

is to emphasise the importance of the statement rather than to quote (Brinton 2008: 84–8), as in (14).

(14) What is that Paper, **I say**? (STEVENS-1745-1,36.443)

The borderline between a reporting clause and a comment clause is not clear-cut so the latter were not easily identifiable. However, the division between reporting clauses and comment clauses was operationalised in the study by means of tense and subject. On the basis of the Penn corpora annotation, the clauses were divided according to tense (past or present verb form) and subject (pronominal or nominal); pronominal subjects were also divided according to person. As shown, for example, by Griffiths (2015), comment clauses are typically in the simple present and have a first-person subject.<sup>5</sup> Therefore, clauses with first-person subjects and present verbs were treated as ambiguous and excluded from the quantitative part of the analysis presented in section 3. Table 1 shows the number of clauses rendered by the queries and finally taken into consideration.

In the analysis presented in section 3, the collocation strength of various reporting verbs and the parenthetical reporting construction is checked. The method applied for this purpose is the collexeme analysis. Its aim is to assess the collocation strength of selected lexical items and the investigated construction, i.e. the degree of mutual attraction between the construction and the lexical element (Stefanowitsch & Gries 2003). The calculations were performed by means of Coll.analysis 3.2a (Gries 2007), a script working in R (2014). In order to calculate collocation strength, it was necessary to establish the overall frequency of all reporting verbs. First, a list of all the verbs which could be used to report speech was prepared for each period on the basis of existing studies (the verbs are listed in relevant fragments of section 3). Next, all finite forms of these verbs were extracted from the corpora by means of the queries such as (VBPI|VBDI|VBP|VBD iDoms

<sup>4</sup> The query \*PRN\* iDoms VB\* also rendered single cases of clause-initial reporting clauses with quotative inversion (see fn. 1), which were excluded at this point.

<sup>5</sup> Griffiths (2015), whose study is based on contemporary data from English, Dutch and German, also suggests that subject–verb inversion may only occur in reporting clauses so the presence of inversion disambiguates the clause. Since in this study inversion is studied independently of function, no such assumptions were made beforehand, but the analysis presented in section 5 largely confirms Griffiths' (2015) finding.

Table 2. *Example of raw data for Coll.analysis (Middle English)*

Verb	Frequency in the corpus	Frequency in the parenthetical reporting construction
<i>quethen</i>	421	373
<i>seien</i>	5,585	993
<i>answeren</i>	307	0
<i>speken</i>	739	0
<i>crien</i>	87	1
<i>callen</i>	235	0

sai\*|say\*|Sai\*|Say\*). Various morphological and spelling variants of each verb in each period were identified in the corpora by means of the lexicon function (make\_lexicon: t pos\_labels: VB\*). Then, a .csv file with the data was prepared for each period (an example of such a dataset is presented in table 2). Finally, the files were fed into Coll.analysis 3.2a, which calculated the collostructional strength for each verb in each period.

Moreover, all the results from sections 3.1–3.4 are summarised in section 3.5 in the form of a decision tree. The data were fed into the Weka implementation (Hall *et al.* 2009) of the C4.5 decision tree algorithm (Quinlan 1993). The tree shows which of the variables included in the data analysis are most useful to model the choice between SV and VS patterns in parenthetical reporting clauses from all stages of English.

Finally, it should be noted that in sections 3.1–3.4 clause-medial and clause-final reporting clauses are presented together without subdivision because the only period in which this factor has any observable impact on the data is Late Modern English (LModE) (as indicated in section 3.5), and the clause-final placement of reporting clauses is generally very rare in all the periods.

### 3 General results

This section presents the general development of the parenthetical reporting construction from Old English, through Middle English and Early Modern English (EModE) to Late Modern English, identifying the main tendencies in inversion rates, showing changes in the use of reporting verbs and the strength of association of particular verbs with the analysed construction.

The normalised frequency of the parenthetical reporting construction is different in each period, as illustrated by figure 1, with a relatively low frequency of occurrence in OE, a sudden peak in ME and a steady decrease through EModE to LModE. This development is to some extent shaped by the composition of the corpora and the growing proportion of non-narrative texts, where quoting speech is by nature less frequent, though it also reflects the change in proportions between clause-initial

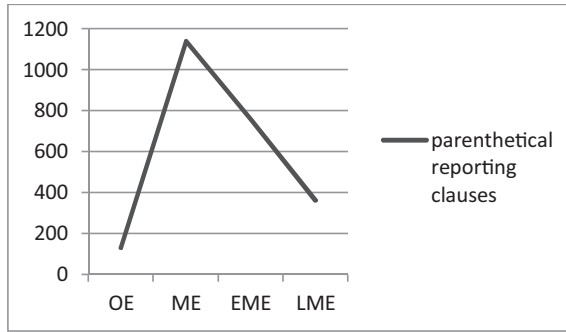


Figure 1. The use of parenthetical reporting clauses in all stages of English (normalised frequencies per 1 million words)

and clause-medial/final placement of reporting clauses in various periods, as explored below.

### 3.1 Old English

As shown in table 3, the whole YCOE corpus (1.5 million words) contains only 196 reporting clauses placed after or in the middle of the reported message, as in (15). None of these clauses shows expansion (i.e. they consist of the subject and the reporting verb only).

- (15) *þæt is soð, cwæð Beotius.*  
 that is true said Boetius  
 ‘That is true, said Boethius’ (coboeth,Bo:26.59.10.1089)

Table 3 shows that inversion always takes place with nominal subjects, as in (15), but there is variation between SV and VS with pronominal subjects, as in (16) and (17).

- (16) *ic eom se þe eom, cwæð he*  
 I am who am said he  
 ‘I am who I am, he said’ (cootest,Exod:3.14.2367)
- (17) *Se ðe eow hyreð, he cwæð, me he gehyreð*  
 this who you hears he said me he hears  
 ‘Who hears you, he said, hears me’ (cowulf,WHom\_17:55.1401)

Table 3. OE parenthetical reporting clauses

Verb	Pronominal subjects			Nominal subjects		
	VS	SV	Total	VS	SV	Total
<i>cweþan</i>	42 (46.1%)	49 (53.9%)	91	98 (100%)	0 (0%)	98
<i>secgan</i>	0 (0%)	3 (100%)	3	4 (100%)	0 (0%)	4
Total	42 (44.7%)	52 (55.3%)	94	102 (100%)	0 (0%)	102

Table 4. *Collostructional strength of cweþan in the parenthetical reporting construction in OE*

Occurrences	<i>cweþan</i>	Other reporting verbs	Total
PRC	189 (2.3%)	7 (0.2%)	196
Elsewhere	8,128 (97.7%)	3,295 (99.8%)	11,423
Total	8,317	3,302	11,619

This variation is restricted to the main OE reporting verb *cweþan* ('say'), which is used in 189 out of 196 occurrences of the parenthetical reporting construction (96 per cent). The use of the other reporting verb (*secgan*, the ancestor of *say*) is very limited (7 instances), and the choice between SV and VS in the case of this verb is clearly based on subject type (nominal ones show inversion, pronominal ones do not).

In order to check whether the high frequency of *cweþan* in this particular construction is a result of its high frequency in the YCOE corpus, a collexeme analysis was performed and the results are presented in table 4. The verbs included in the calculation as 'other reporting verbs' are: *secgan* ('say'), *andswarian* ('answer'), *frignan* ('ask') and *ascian* ('ask').

It turns out that *cweþan* is the only OE reporting verb attracted to the PRC; despite its overall high frequency, *cweþan* appears in the construction more frequently than expected (attraction, CollStr=19.15,  $p < 0.001$ ). However, direct speech in OE tended to be introduced by reporting clauses placed before a quotation, and the parenthetical construction was a minority pattern. When *cweþan* (or any other reporting verb) introduced direct speech in OE, clause-initial placement of the reporting clause, as in (18), was most common.

- (18) **Heo him andswarode:** gehyr nu  
 she him answered hear now  
 'She answered him: Hear now' (cogregdC,GD\_2\_[C]:33.168.10.2061)

### 3.2 Middle English

As signalled by figure 1, in ME parenthetical reporting clauses become relatively numerous (1,367 instances). However, table 5 shows that the inversion rates of pronominal and nominal subjects are close to the ones identified for OE, though the behaviour of subject pronouns in the ME data depends on the reporting verb.

There are two verbs used regularly in the investigated construction: *quethen* and *seien*. The verb *quethen*, which is a direct descendant of OE *cweþan*, regularly inverts all subjects regardless of their type, as in (19) and (20).

- (19) 'Nay!' **quod Thomas,** 'no traytour, but þe archbyschop.'  
 no said Thomas no traitor but the archbishop  
 "No," said Thomas, "no traitor but the archbishop" (CMMIRK-M34,42.1212)



Table 5. *ME parenthetical reporting clauses*

Verb	Pronominal subjects			Nominal subjects		
	VS	SV	Total	VS	SV	Total
<i>quethen</i>	283 (99.6%)	1 (0.4%)	284	89 (100%)	0 (0%)	89
<i>seien</i>	84 (24.3%)	261 (75.7%)	345	641 (98.9%)	7 (1.1%)	648
<i>crien</i>	0 (–)	0 (–)	0	1 (100%)	0 (0%)	1
Total	367 (58.3%)	262 (41.7%)	629	731 (99.1%)	7 (0.9%)	738

Table 6. *Collostructional strength of quethen and seien in the parenthetical reporting construction in ME*

Frequency	<i>quethen</i>			<i>seien</i>		
	<i>quethen</i>	Other r. verbs	Total	<i>Seien</i>	Other r. verbs	Total
PRC	373 (89.0%)	994 (14.3%)	1,367	993 (17.8%)	374 (20.9%)	1,367
Elsewhere	46 (11.0%)	5,961 (85.7%)	6,007	4,592 (82.2%)	1,415 (79.1%)	6,007
Total	419	6,955	7,374	5,585	1,789	7,374

- (20) That is sooth, **quod** I  
 that is true said I  
 ‘That is true, I said’ (CMBOETH–M3,434.C1.230)

In the case of *seien*, derived from OE *secgan*, variation between SV and VS may be observed. Nominal subjects predominantly undergo inversion, as in (21), while non-inversion prevails with pronominal subjects (76 per cent), as in (22). (21) also shows that clauses with expansion regularly co-occur with the VS order (unlike in PDE).

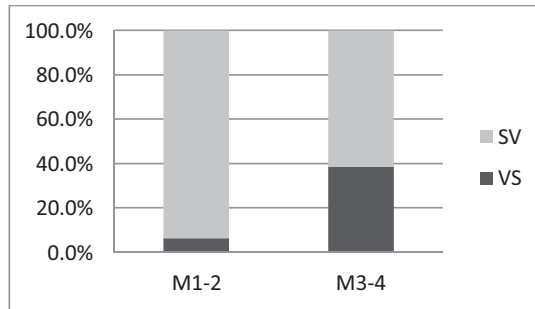
- (21) ‘Sir,’ **seyde** **Merlion** **unto** **the** **kynge**, ‘woll ye geff me  
 sir said Merlin unto the king will you give me  
 a gyffte?’  
 a gift  
 ‘“Sir,” said Merlin to the king, “will you give me a gift?”’ (CMMALORY–M4,30.937)
- (22) ‘Lord!’ **he** **seyde**, ‘come down byfore þat my sone dye.’  
 Lord he said come down before that my son die-SBJ  
 ‘“Lord,” he said, “come down before my son dies.”’ (CMWYCSE–M3,305.1405)

In order to check the strength of association between *quethen* and the parenthetical reporting construction, and to investigate whether *seien* became attracted to the construction, a collexeme analysis was performed (table 6). Other reporting verbs included in the calculation were: *answren*, *speken*, *crien*, *callen* (after Moore 2015: 258).

The analysis reveals that *quethen* is strongly attracted to the construction (attraction, Coll.Strength = 232.78,  $p < 0.001$ ), but *seien* is not (repulsion, Coll.Strength = 2.73,  $p < 0.01$ ), even though parenthetical reporting clauses with *seien* outnumber those

Table 7. *ME parenthetical reporting clauses (excluding Malory)*

Verb	Pronominal subjects			Nominal subjects		
	VS	SV	Total	VS	SV	Total
<i>quethen</i>	283 (99.6%)	1 (0.4%)	284	89 (100%)	0 (0%)	89
<i>seien</i>	33 (12.3%)	234 (87.7%)	267	148 (98.0%)	3 (2.0%)	151
Total	316 (57.3%)	235 (42.7%)	551	237 (98.7%)	3 (1.3%)	240

Figure 2. Reporting clauses with *seien* and pronominal subjects in ME<sup>6</sup>

containing *quethen* (993 vs 373). The dataset is dominated by *seien* (76 per cent share of all the investigated verbs), which replaced *quethen* as the most frequent and diverse reporting verb, used in a variety of contexts, the reporting parenthetical construction being just one of them, see (23).

- (23) **He seith also**, ‘Ther may no thyng be likned to the trewe freend’  
 he says also there may no thing be likened to the true friend  
 ‘He also said: “Nothing may be compared to a true friend”’ (CMCTMELI,223.C1.228)

However, the proportions between *seien* and *quethen* are to some extent shaped by intertextual differences: as many as 576 parenthetical reporting clauses extracted from the corpus come from a single text from late ME, namely Malory’s *Le Morte D’Arthur*; all of them (except a single occurrence of *crien*) contain the verb *seien*. Table 7 shows what happens to the ME data if *Le Morte D’Arthur* is taken out of the study sample. It turns out that the tendency towards non-inversion of pronominal subjects with *seien* is further strengthened, but the overall tendencies are confirmed. However, the fact that inversion of subject pronouns with *seien* is more common in Malory (51 out of 78 subject pronouns are inverted, 65 per cent of VS) than in other texts from PPCME2 (12 per cent of VS) is a reflection of a diachronic trend, illustrated in figure 2.

<sup>6</sup> For the diachronic analysis, the periodisation from the Helsinki corpus was treated as the basis for the division into early ME (M1, M2) and late ME (M3, M4). The subperiods were combined because of the differences in word count (M2 has only 93,999 words) and due to the fact that Malory completely dominates the M4 subperiod. If the period of composition and the period of the manuscript were different, the former was treated as decisive. *Le Morte D’Arthur* is classified in the corpus as M4.

As shown in the figure, the less frequent inverted pattern, which is hardly present in early ME, reaches almost 40 per cent in late ME; the change is statistically significant (chi-square = 48.53, p-value < 0.0001) and it shows that the tendency to invert subject pronouns with *seien* is strengthened during the ME period.

### 3.3 Early Modern English

In EModE, illustrated in table 8, nominal subjects are still regularly inverted (like in OE and ME) and the overall rate of inversion of pronominal subjects rises to 89 per cent (from 58 per cent in the previous period), which means that the diachronic trend identified in the ME corpus is continued. (24) is an example of the dominant VS pattern.

(24) but pray, **says he**, do not tell my Mistress of it (LISLE–E3–P2,4.119.414)

Non-inversion of nominal subjects is still very rare and observed only with *say*, as in (25); inversion illustrated with (26) is predominant.

(25) But only he would be mayster of his horsse, **the Scripture sayeth**  
(LATIMER–E1–H,32L.240)

(26) ‘Alas!’ **sayth this good wyfe**, ‘he is to stronge for you all.’ (HARMAN–E1–H,71.293)

Table 8 shows that *say* retains its position as the dominant reporting verb, and the frequency of the non-inflected form *quoth*, which according to the *Oxford English Dictionary* became a fixed quotation marker used for all persons and only in its past form by the end of the sixteenth century, is still very high; VS shown in (27) is the only pattern co-occurring with this verb. As can be seen, specification of the addressee has no impact on inversion, which is obligatory.

Table 8. EModE parenthetical clauses with verbs of saying

Verb	Pronominal subjects			Nominal subjects		
	VS	SV	Total	VS	SV	Total
<i>say</i>	311 (79.5%)	80 (20.5%)	391	252 (96.5%)	9 (3.5%)	261
<i>quoth</i>	360 (100%)	0 (0%)	360	253 (100%)	0 (0%)	253
<i>reply</i>	4 (100%)	0 (0%)	4	5 (100%)	0 (0%)	5
<i>tell</i>	0 (0%)	4 (100%)	4	0 (–)	0 (–)	0
<i>answer</i>	5 (100%)	0 (0%)	5	0 (–)	0 (–)	0
<i>call</i>	2 (100%)	0 (0%)	2	0 (–)	0 (–)	0
<i>return</i>	2 (100%)	0 (0%)	2	0 (–)	0 (–)	0
<i>add</i>	1 (100%)	0 (0%)	1	1 (100%)	0 (0%)	1
<i>continue</i>	1 (100%)	0 (0%)	1	0 (–)	0 (–)	0
<i>cry</i>	0 (–)	0 (–)	0	1 (100%)	0 (0%)	1
Total	686 (89.1%)	84 (10.9%)	770	512 (98.3%)	9 (1.7%)	521

Table 9. *Collostructional strength of quoth and say in the parenthetical reporting construction in EModE*

Frequency	<i>quoth</i>			<i>say</i>		
	<i>quoth</i>	Other r. verbs <sup>7</sup>	Total	<i>say</i>	Other r. verbs	Total
PRC	613 (80.4%)	678 (6.8%)	1,291	652 (13.2%)	639 (11.1%)	1,291
Elsewhere	149 (19.6%)	9,225 (93.2%)	9,374	4,281 (86.8%)	5,093 (88.9%)	9,374
Total	762	9,903	10,665	4,933	5,732	10,665

(27) And what is your name **quoth the Constable to Meg?** (PENNY-E3-P1,52.327)

What is more, some new reporting verbs appear in the corpus; examples are shown in (28)–(30).

(28) But in good faith **added the Cobler** I am resolved to be merry with you. . .  
(PENNY-E3-P1,28.108)

(29) ‘So he was,’ **answered I** (RICH-E2-P1,1.3,172.23)

(30) I well perceived that, **returned she**, when with silence and attention thou didst receive my Words (BOETHPR-E3-P2,92.9)

The statistical analysis presented in table 9 shows that *quoth* is extremely attracted to the construction (attraction, Coll.Strength is infinite,  $p < 0.001$ ) since it appears in the corpus mainly as a part of it, which constitutes a significant difference between *quoth* and other reporting verbs (80 vs 7 per cent). The strength of association between *say* and the construction is lower (attraction, Coll.Strength=3.21,  $p < 0.001$ ), but unlike in OE and ME the verb is attracted to the construction. None of the other reporting verbs included in the analyses shows a strong association with the parenthetical construction.

### 3.4 Late Modern English

In LModE, compared to EModE, the frequency of quotative inversion falls both for nominal and pronominal subjects, though the inversion rate is still relatively high for the former, as shown in table 10. The dominant reporting verb is *say*, which shows variation between VS and SV with pronominal subjects, as in (31) and (32), while nominal subjects are mostly inverted as in (33).

(31) ‘It had worn him down,’ **he said** (SOUTHEY-1813-1,184.202)

(32) ‘O,’ **said he**, ‘how much I would give to hear some of your private conversations!’  
(BURNEY-1768-2,1,30.665)

(33) my dear youth, **said sir Philip**, no apology is necessary (REEVE-1777-1,13.312)

<sup>7</sup> All the reporting verbs listed in table 8 are included.

Table 10. *LModE parenthetical clauses with verbs of saying*

Verb	Pronominal subjects			Nominal subjects		
	VS	SV	Total	VS	SV	Total
<i>say</i>	251 (67.1%)	123 (32.9%)	374	270 (92.8%)	21 (7.2%)	291
<i>cry</i>	32 (71.1%)	13 (28.9%)	45	26 (100%)	0 (0%)	26
<i>continue</i>	37 (86.0%)	6 (14.0%)	43	10 (90.9%)	1 (9.1%)	11
<i>add</i>	33 (82.5%)	7 (17.5%)	40	2 (100%)	0 (0%)	2
<i>reply</i>	30 (83.3%)	6 (16.7%)	36	7 (100%)	0 (0%)	7
<i>answer</i>	16 (72.2%)	6 (27.3%)	22	15 (100%)	0 (0%)	15
<i>tell</i>	0 (0%)	9 (100%)	9	0 (0%)	4 (100%)	4
<i>write</i>	0 (0%)	8 (100%)	8	12 (92.3%)	1 (7.7%)	13
<i>quoth</i>	12 (100%)	0 (0%)	12	0 (-)	0 (-)	0
<i>return</i>	5 (100%)	0 (0%)	5	7 (100%)	0 (0%)	7
<i>ask</i>	0 (0%)	2 (100%)	2	0 (0%)	4 (100%)	4
<i>exclaim</i>	1 (33.3%)	2 (66.7%)	3	1 (100%)	0 (0%)	1
<i>repeat</i>	1 (100%)	0 (0%)	1	2 (100%)	0 (0%)	2
<i>go on</i>	0 (0%)	1 (100%)	1	0 (0%)	3 (100%)	3
<i>remark</i>	0 (0%)	1 (100%)	1	2 (100%)	0 (0%)	2
<i>resume</i>	0 (0%)	1 (100%)	1	2 (100%)	0 (0%)	2
other	0 (0%)	10 (100%)	10	11 (73.3%)	4 (16.7%)	15
Total	418 (68.2%)	196 (31.8%)	613	367 (90.6%)	38 (9.4%)	405

Another change compared to the earlier period is the repertoire of reporting verbs, which rises substantially in the LModE period: *quoth* is hardly ever used (only 12 occurrences, all of them with pronominal subjects), but there are as many as 38 various reporting verbs attested in the corpus, though only some of them have a relatively high frequency, as shown in table 10. The more frequent verbs show a tendency for inversion of all subject types, e.g. *cry* in (34), *add* in (35) and *reply* in (36), while the less frequent ones rarely cause inversion of subject pronouns.

- (34) 'I wish the ladies would put themselves under my care,' **cried Morrice**, 'and take a turn round the park.' (BURNEY-1782-2,1,139.575)  
 (35) And, **adds he**, Tho' I hate his Principles, yet I would not have him fall into their Hands (DEFOE-1719-1,199.109)  
 (36) It was never my intention, **replied she**, to entertain you with delusions. (BOETHRI-1785-2,131.903)

The strength of association between *say* and the analysed construction is illustrated by table 11. While other reporting verbs appear in the parenthetical reporting construction in only 4 per cent of cases, *say* co-occurs with it in 16 per cent of its uses. However, new verbs are also attracted to the construction, as illustrated in table 12. The one which is most strongly associated with the analysed structure is *say* (attraction, Coll.Str=115.22,  $p < 0.001$ ), but *cry*, *reply* and *add* also start to co-occur with it with a significant frequency.

Table 11. *Collostructional strength of say in the parenthetical reporting construction in LModE*

Occurrences	<i>say</i>	Other reporting verbs	Total
PRC	665 (15.9%)	353 (4.2%)	1,018
Elsewhere	3,510 (84.1%)	8,074 (95.8%)	11,584
Total	4,175	8,427	12,602

Table 12. *Reporting verbs attracted to the parenthetical reporting construction in LModE*

Reporting verb	Parenthetical uses	Total frequency	CollStr
<i>say</i>	665	4,175	115.22
<i>cry</i>	71	317	15.62
<i>quoth</i>	12	18	9.21
<i>reply</i>	43	285	4.62
<i>add</i>	42	413	1.31

### 3.5 Summary

The general diachronic analysis reveals a few interesting developments. First of all, there is a steady increase in the association between the parenthetical reporting construction and two main reporting verbs: *say* and *quoth*. Moreover, the corpus analysis shows a growing type-frequency of verbs co-occurring with the construction (from two in OE to 38 in LModE) and verbs attracted to the construction (from a single verb in OE to five verbs in LModE), as well as a sudden increase in token-frequency in the ME period (see figure 1). Next, figure 3 shows that quotative inversion with nominal subjects is the norm in all the historical periods, though the frequency of the inverted pattern starts to fall in the LModE period.

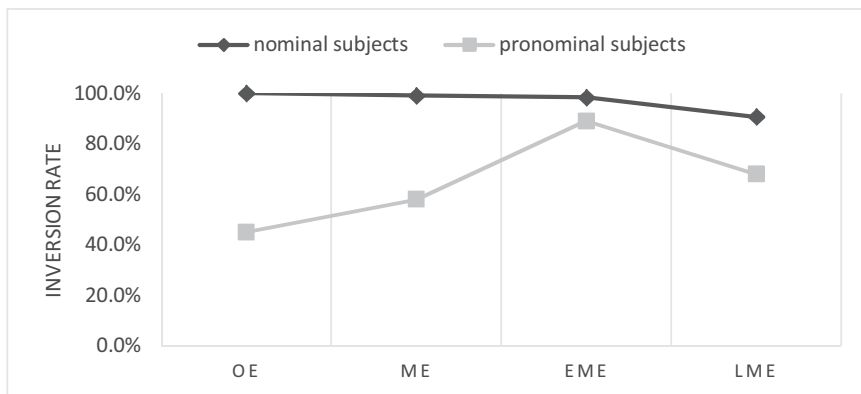


Figure 3. Frequency of the VS pattern in reporting parentheticals from OE to LModE

What is more, variation between the inverted and the non-inverted pattern with pronominal subjects is attested in all the periods, though the development is not regular because the frequency of inversion shows a steady growth from OE to EModE and then falls in LModE, which suggests that it could have been influenced by a variety of factors. In order to highlight the variables which have the strongest impact on the choice between SV and VS in reporting parenthetical clauses, a decision tree has been generated for the whole diachronic dataset.

Subject type = nominal: VS (1763/50)

Subject type = pronominal

| Verb = *quoth*

|| Period = OE: SV (91/42)

|| Period = ME: VS (285/1)

|| Period = EModE: VS (360)

|| Period = LModE: VS (12)

| Verb = *say*

|| Period = OE: SV (2)

|| Period = ME: SV (359/85)

|| Period = EModE: VS (391/79)

|| Period = LModE

||| Position = medial: VS (350/102)

||| Position = final: SV (24/3)

| Verb = other

|| Expansion = no: VS (205/49)

|| Expansion = yes: SV (42/14)<sup>8</sup>

Quite unsurprisingly, it turns out that the main variable underlying the variation is subject type (which causes the primary split), with nominal subjects strongly preferring the VS pattern in all the periods. Within the set of clauses with pronominal subjects, the tree splits between particular verbs, with *quoth* favouring the inverted pattern in all the periods except OE, and *say* moving from SV in OE and ME, through VS in EModE and a position-based variation between SV and VS in LModE (clause-medial reporting clauses showing more VS, which confirms Quirk *et al.*'s observation for PDE (1985: 1022), and clause-final ones being more prone to SV, though their frequency is very low). Other reporting verbs are the only group in which expansion of the clause has an impact on inversion, leading to more regular SV with subject pronouns. There

<sup>8</sup> Correctly classified instances: 88.62 per cent; SV: precision 0.719, recall 0.534; VS: precision 0.91, recall 0.958. Though the overall success rate is quite high, the model is much more successful in predicting the dominant VS variant.

is no split into periods in the case of ‘other’ verbs, but it should be borne in mind that *quoth* and *say* are the only reporting verbs used in OE and ME, and they dominate in EModE, so this result is of importance for LModE only.

All in all, it turns out that the identified patterns should be analysed for each reporting verb separately since they show very distinct behaviour in the historical data, especially with respect to pronominal subjects. Hence, in order to understand the mechanisms behind the identified trends, the development of *quoth* and *say* is analysed in sections 4 and 5 respectively, while the overall changes in the analysed construction are summarised and presented in section 6.

#### 4 The story of *quoth*, the oldest English parenthetical reporting verb

As shown in the previous section, OE reporting parentheticals predominantly contain the verb *cwēpan* (the ancestor of *quoth*). It is found in 96 per cent of clauses with nominal subjects (where VS is the only attested pattern) and it is the only verb co-occurring with inverted subject pronouns, though SV is also noted (see table 3). Harbert reports that in OE ‘pronominal subjects always followed the verb when the latter was the quotative verb *cwēpan* and the sentence-initial position was occupied by (all or part of) its quotative complement’ (2007: 414). His analysis, however, is based only on Bede’s *Historia Ecclesiastica* where the pattern *he cwæð* is never used. A more comprehensive analysis reveals that *cwēpan* shows regular quotative inversion but only with nominal subjects, while subject pronouns exhibit text-based variation between VS and SV. As noticed by Mitchell (1985: §1949), there are two competing patterns of parenthetical insertion with subject pronouns in OE: *cwæð he* used in Bede, as in (37), and *he cwæð* used in the Homilies of Wulfstan, as in (38).

(37) Eala broðor Ecgberht; eala, **cwæð he**, hwæt dydest þu?  
oh brother Egbert oh said he what did you  
‘Oh, brother Egbert, said he, what did you do?’ (cobede,Bede\_3:19.244.3.2498)

(38) Se ðe eow hyreð, **he cwæð**, me he gehyreð  
this who you hears he said me he hears  
‘Who hears you, he said, hears me’ (cowulf,WHom\_17:55.1401)

This study shows that the variation is not limited to these two texts because both patterns are attested in other sources as well, with (39) and (40) as relevant examples. It turns out that 36 out of 49 non-inverted subject pronouns (69 per cent) come from Wulfstan’s texts, while 20 out of 42 inverted ones (48 per cent) are from Bede. Moreover, both texts are consistent because there is not a single instance of SV in Bede or VS in Wulfstan.

(39) Bidað ge her, **he cwæð**  
wait-IMP you here he said  
‘Wait here, he said’ (coverhomE,HomS\_24.1\_[Scragg]:65.31)

(40) ic eom se þe eom, **cwæð he**  
I am who am said he  
‘I am who I am, he said’ (cootest,Exod:3.14.2367)



A similar consistency (though with a lower number of instances) may be observed in Euphrosine (VS), Margaret (C) (VS), Orosius (VS), the Heptateuch (VS), as well as Cura Pastoralis (SV) and the Blickling Homilies (SV).<sup>9</sup> What is more, individual preference may also be observed in the use of *cweþan* with nominal subjects. Out of 98 occurrences of the structure, 47 (48 per cent) come from the OE version of *Historia adversus paganos* by Orosius, and 46 of these have exactly the same, lexically recurrent form *cwæð Orosius* illustrated in (41).<sup>10</sup> This, however, has no impact on inversion rates; if a reporting clause with a nominal subject is used parenthetically, inversion is the only available option.

- (41) ure ieldran ealne þisne ymbhwyrft þises middangeardes,  
 our elders all this ring this-GEN world-GEN  
**cwæþ Orosius,** swa swa Oceanus utan ymbligeþ,  
 said Orosius as Ocean out surrounds  
 þone man garsæg hateð, on þreo todældon  
 which man spear-man calls in three divided  
 ‘Our ancestors the border of the whole world, said Orosius, as it is surrounded by the  
 ocean, called spear-man [Neptune], divided in three’ (coorosiu,Or\_1:1.8.11.101)

Moreover, it is quite remarkable that while according to the *OED* *quoth* became an invariant quotative marker used for all persons and with no present time reference in EModE, its use in the present tense is already limited in the OE period. There are only two parenthetical uses of *cweþan* in the present form and they both co-occur with the first-person pronoun *ic* (‘I’) and follow the SV pattern, as in (42) (they were excluded from the study sample according to the methodology described in [section 2](#)).

- (42) Oþer is, **ic cweþe,** se æresta apostol, oþer se nehsta;  
 other is I say the first apostle other the next  
 ‘One is, I say, the first apostle, and the other the next’  
 (coblick,LS\_32\_[PeterandPaul[BiHom\_15]]:171.5.2158)

The example resembles a comment clause rather than a typical reporting structure, but such a use of *cweþan* in OE is clearly exceptional. What is more, among the 189 occurrences of *cweþan* in the parenthetical reporting construction (which are all in the past tense), only one, shown in (43), has a plural ending.

- (43) Ge Galileiscan weras, **cwædon hie,** forþon þe hie  
 you Galilean men said they because they  
 wæron of Galileam þæm lande, hwæt stondað ge her  
 were of Galilee the land what stand you here

<sup>9</sup> Other texts contain only single instances of either pattern, and Vercelli Homilies are the only text in which single occurrences of both *cwæð he* and *he cwæð* are attested, though since it is a collection of homilies of unknown (and possibly collective) authorship, lack of consistency is not surprising.

<sup>10</sup> Orosius is a translation but the pattern shown in (41) is not influenced by the original Latin text since the whole phrase *cwæð Orosius* was added by the translator without any clear model in the source text (which is logical considering the fact that Orosius was the author of the Latin original).

& pyses wundriað, & up on pysne heofon lociaþ?  
 and this wonder and up on this heaven look  
 ‘You, Galilean men, they said, because they were from the land of Galilee, why  
 are you standing here and wondering about this and looking up to heaven?’  
 (coblick,HomS\_46\_[BIHom\_11]:123.131.1547)

This means that in 99 per cent of the instances, the verb has the fixed past form *cwæþ*, used in the first and third person singular according to the rules of OE grammar, as in (44) and (45).

(44) Gea **cwæð** he  
 yes said he  
 ‘Yes, he said’ (coeuphr,LS\_7\_[Euphr]:67.64)

(45) Gyse, **cwæð** ic  
 yes said I  
 ‘Yes, I said’ (coboeth,Bo:34.87.12.1666)

Therefore, it is plausible to assume that the process of constructionalisation of *quoth* as an invariant quotation marker started already in OE, where the morphologically fixed form was dominant, though it had not spread to the plural paradigm yet.

In the ME period, *quethen* becomes a part of a syntactically fixed VS quotative construction used for both subject types, as shown in table 7. What is more, already in ME it is used only in its past form *quod/quob* and among 373 occurrences of the verb in the investigated structure, only two as in (46) have the plural ending *-en* and only one other shown in (47) does not follow the VS pattern. (None of the analysed examples is interpretable as a comment clause; their function is clearly to introduce quotations.)

(46) Nu **cweden** ha. wa him þe ne fondi to-dei  
 now said they woe him that not found today  
 for te wurche þe wurst.  
 for to work the worst  
 ‘Now, they said, woe to him who did not find today the worst to work.’  
 (CMJULIA–M1,104.136)

(47) Lauerd **he cweð** þa. Nu ic þe bidde for þine kinedome...  
 lord he said then now I you ask for your kingdom  
 ‘Lord, he said then, now I ask you for your kingdom’ (CMLAMBX1–MX1,45.585)

These untypical examples come from the earliest part of the corpus (M1), so their presence is most likely a remnant of the patterns existing in the OE period. The fact that the plural form is used only twice does not mean, however, that the subjects of reporting clauses in ME are all singular, see (48).

- (48) A, lady, **quod they**, ‘ye han shewed unto us the blessing  
 Oh lady said they you have showed unto us the blessing  
 of swetnesse, after the sawe of David the prophete  
 of sweetness after the example of David the prophet  
 ‘Oh lady, they said, you have shown us the blessing of sweetness like David the  
 prophet’ (CMCTMELI–M3,236.C2.762)

The presence of such clauses (4 instances of *quod they*, all from the M3 period) shows that *quod/quoth* became a morphologically fixed quotation marker irrespective of person and number already in the ME period, so earlier than the *OED* suggests.

What is more, in the ME period the verb *quethen*, or rather its most frequent form *quod/quoth*, undergoes spelling reduction and appears in its abbreviated form *qð* as in (49) 76 times (20 per cent). As suggested by Moore (2015: 260), the reduction of *quod* to a symbol looking like a punctuation mark rather than a fully fledged word is analogous to the phonetic reduction observed in speech when grammaticalisation takes place.

- (49) Nai **qð ha**  
 no said he  
 ‘No, he said’ (CMMARGA–M1,59.68)

In the EModE period, the tendencies identified in the ME data are continued. The verb appears only in the past tense and without plural marking, as in (50). Inversion is obligatory but spelling reduction is visible in only 32 examples (out of 613, 5 per cent),<sup>11</sup> as illustrated by (51).

- (50) Why, haue you no more? **quoth they** (HARMAN–E1–P2,54.328)  
 (51) Yes sir **qd. Simon** if she be as willing as I (PENNY–E3–P1,127.569)

In LModE, the frequency of *quoth* falls drastically and the verb becomes obsolete, thus giving way to new reporting verbs, which start to appear in LModE with a growing frequency.

## 5 The development of *say*-parentheticals

In PDE, *say* is the most frequent reporting verb, but in early English its position was much weaker. As shown in table 3, the frequency of its OE ancestor *secgan* (‘say’) in parenthetical reporting clauses is very low. In ME, the frequency of *seien* soars. It becomes the main reporting verb and its use in the parenthetical construction grows (though section 3.2 shows that it is not attracted to the construction yet). Section 4 demonstrates that at that time *quod/quoth* was already a fixed quotation marker used

<sup>11</sup> Moore (2015) associates the reduction observed in ME to inconsistent marking of quotations in ME manuscripts. The reduced form *qd* or even *q* served as a quotation mark, so the decrease in the use of the reduced variant in EModE may be interpreted as a result of growing standardisation of punctuation marking after the introduction of print.

Table 13. *Difference in inversion rates between two groups of say-parentheticals*

Period	<i>say</i> -parentheticals with first-person subjects and present verbs			Other <i>say</i> -parentheticals with pronominal subjects		
	SV	VS	Total	SV	VS	Total
EModE	96 (92.3%)	8 (7.7%)	104	80 (20.5%)	311 (79.5%)	391
LModE	110 (93.2%)	8 (6.8%)	118	123 (32.9%)	251 (67.1%)	374

only in the past tense and without plural marking, so it is plausible that there was a need for a new reporting verb which would be more flexible and show a full set of grammatical properties (number, person and tense). It turns out that *c.* 26 per cent of reporting clauses with *say* in ME (255 out of 993) are in the present tense. Thus, the fact that *quethen* could no longer appear in certain grammatical contexts may have opened up the floor for *seien*, which took up the functions that *quethen* could not perform and gradually started to push *quethen* out of the system.

The inverting influence of *seien* is weaker compared to *quethen*, especially in early ME, as shown in section 3.2, though the difference is hardly visible in the case of nominal subjects. In late ME, subject pronouns start to invert with *say* more regularly, and in EModE the VS pattern becomes dominant. It may be considered surprising since this development goes against the PDE tendency for non-inversion of pronominal subjects in parenthetical clauses. There are two plausible (and possibly complementary) reasons for this unexpected change: (i) analogy, with the high frequency of *quod/quob* parentheticals in ME functioning as a general model for the VS pattern used in reporting clauses irrespective of subject type, and (ii) the development of a competing parenthetical construction, namely a comment clause with *say*. The earliest (rare) examples of this structure are found already in ME, illustrated in (52), but it does not become relatively frequent until EModE, as in (53).<sup>12</sup>

(52) Thus, **I say**, Thomas seruet God deuotly (CMMIRK–M34,41.1176)

(53) Therefore take heede, **I saye**, for Christes sake (THROCKM–E1–H,I,76.C2.727)

The function of comment clauses is to emphasise the importance of the statement (Brinton 2008: 84–8; Traugott 2012: 471) rather than to report someone's words or thoughts. Table 13 shows that *say*-clauses with a potentially commenting function show a clear tendency for the SV order, as in (52)–(53).

Thus, it is evident that if *say* was used with a reporting function, inversion was the preferred pattern, and SV was the default order of the comment clause (which is still used today, see Griffiths 2015). This development may be traced back to EModE, when the number of comment clauses with *say* started to grow, and there was an overlap

<sup>12</sup>As explained in section 2, such clauses were excluded from the quantitative part of the study and they are not included in the data presented in section 3.

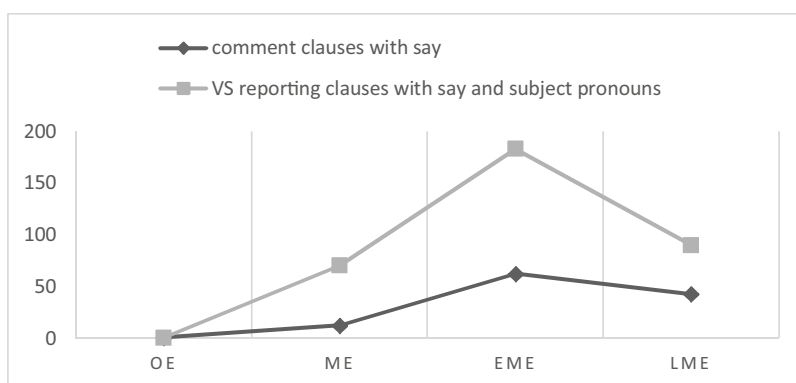


Figure 4. The development of *say* in comment clauses and VS parenthetical reporting clauses with subject pronouns (normalised frequencies per 1 million words)

between reporting and commenting parentheticals with *say*. Figure 4 shows that these two developments co-occur.

A similar tendency may be observed for the variation between *he says/said vs says/said he*. In the latter case, the clauses are all clearly reporting, as in (54). In the former, however, they are often parentheticals inserted in the middle of a long fragment which is a story narrated from someone's perspective in a free indirect style rather than a direct quote as in (55).

(54) Are you my uncle? **says hee againe**. (ARMIN–E2–H,43.299)

(55) After this, master Lieutenant, cominge into his chamber to visite him, rehearsed the benefittes and freindshipp that he had many waies receaved at his handes, and howe much bounden he was therefore freindly to intertayne him, and make him good cheare; which, since the case standing as it did, he could not do without the kinges indignation, he trusted, **he said**, he wold accepte his good will, and suche poore cheare as he had. (ROPER–E1–P2,77.83)

Thus, it seems that in EModE inversion started to be a form strongly associated with quoting speech,<sup>13</sup> and the non-inverted patterns may be interpreted as non-reporting (mostly comment) clauses which are more loosely connected to the content of the main clause. It transpires that it is function, and not subject type, which was the factor underlying the choice between VS and SV in parenthetical clauses with *say* in EModE.

Another interesting phenomenon related to the development of *say*-parentheticals is the emergence of the invariant singular present form *says I/he*. As shown in table 13, among the clauses excluded from the study sample (with present tense verbs and first-person subjects), there are some isolated cases of VS. In the EModE data there are

<sup>13</sup>As pointed out by one of the reviewers, a similarly unexpected increase in inversion rates for subject pronouns in ME is also noted in other syntactic contexts by van Kemenade (1987: 198) and Haerberli (2002: 259ff.) in their studies of the loss of V-2 in English. It is beyond the scope of this study to assess to what extent these two developments are interrelated, but it may be another argument confirming that the V-2 rule and quotative inversion interacted in the ME period, as explored in detail in section 6 of this article.

five<sup>14</sup> instances of the sequence *say I*, which is not attested in PDE (Biber *et al.* 1999: 922); there are no examples of the non-standard form *says I* in this period. Four out of these five uses come from the same fragment of the same text, as shown in (56). The clause *say I* appears in a sequence of quoted utterances, and its function is clearly reporting.

- (56) Who is this (**sayth one**) sir Launcelot du lake?  
 Who is this, greate Guy of Warwike, **sayth an other**;  
 No (**say I**) it is the thirteenth Hercules brother.  
 Who is this; noble Hector of Troy, **sayth the thirde**;  
 No, but of the same nest (**say I**) it is a birde.  
 Who is this; greate Goliah, Sampson, or Colbrande?  
 No (**say I**) but it is a brute of the Alie lande.  
 Who is this; greate Alexander? Or Charle le Maigne?  
 No, it is the tenth Worthie, **say I to them agayne**:  
 I knowe not if I sayd well. (UDALL–E1–H,L.165.61-71)

The form *says/saith I* appears for the first time in the LModE corpus with five occurrences of the pattern illustrated in (57), though the form *say I* shown in (58) is also attested. What is more, the *-s* ending is also used with the pronoun *you*, as in (59).

- (57) Mr. Layer he look'd at me, he was really a perfect Stranger to me, for I did not know him again; saith he, Is not your Name Plunkett? Yes, **saith I**. Was there not one with you t'other Night, one James Plunkett? Yes, Sir, **says I**. Where did he desire you to go? said he. I answer'd, He desired me to go to the Italian Coffee-house in Russel-Court. Saith he, 'Tis well enough: Do you not know me? No, **saith I** (LAYER–1723–2,57.2321-2332)
- (58) Hastings: Damn your pig, I say. Marlow: And damn your prune sauce, **say I**. (GOLDSMITH–1773–1,28.255-256)
- (59) On a sultry day, when both the sun and the enemy had set us in a glow, your groom was milling with your canteen; you came to me – Warmans, **says you**, have you any thing to drink? – I reached my flask – you drank – did you not? (JOHNSTONE–1786–2,44.744-748)

The examples in (57)–(58) were excluded from the analysis presented in section 3. However, the function of all the instances of *saith/says I* shown in (57) is clearly reporting, and so is *says you* in (59), while (58) functions as a comment clause (an expression of emphasis), used facetiously by Marlow as a reaction to Hastings' *I say* (and with contrastive focus on the inverted subject). It is difficult to make generalisations on the basis of such a low number of instances, but it seems that in LModE inversion or its lack was not a clear enough marker of function and the invariant pattern *says + pronoun* started to be generalised as a LModE reporting formula. The reason for its emergence may have been phonological (to avoid hiatus

<sup>14</sup>Table 13 shows that there are eight VS *say*-parentheticals in EModE: five of them have the form *say I*, while the remaining three clauses have a complex verb phrase consisting of a modal verb followed by *say*, e.g. *would I say*.

between two diphthongs in *say I*). However, it could also be viewed as the beginning of a morphological reduction. The fact that singular third person subjects are predominant in reporting clauses made *says/saith* a majority pattern and the fact that clauses such as *says/saith I* start to appear in LModE suggests that *say* may have been on its way to undergo a similar grammaticalisation process as *quoth* (though with two separate tense forms: *said* and *says/saith*). This process was most probably hindered by prescriptive pressure, but since *says I* is reported to be used in PDE conversations (Biber *et al.* 1999: 922), it may have survived in spoken English, though (as mentioned in section 1) apparently with an archaic flavour. The few examples of the pattern that I was able to extract from COCA<sup>15</sup> represent speech in written sources, i.e. dialogues in fiction as in (60).

- (60) It's your plant,' **says I**. I gave it to you.  
(COCA, 1997, *Literary Review*, Philip Davidson)

What is more, COCA contains 675 instances of the string *I says*, as in (61), and in the spoken component of the BNC there are 908 occurrences, as in (62), which has both patterns used within the same utterance. This shows that in PDE the third-person ending may be used as a quotation marker regardless of inversion.<sup>16</sup> Thus, for some of the LModE and PDE data it is not really inversion which is decisive for the functional interpretation of the clause (as it is in EModE), but the use of the ending, generalised to all persons and numbers, as illustrated by (63).

- (61) 'Yes, Jack,' **I says** (COCA, 2001, *The Death of Jack Hamilton*, Stephen King)  
(62) Oh, **says I**, aye, **I says**, you could stay in the village. (Spokes BNC, Oral history project: interview)  
(63) Won't tolerate babblin', **they says**. (COCA, 1994, *Thief of the Hearts*, Teresa Medeiros)

The fact that *I says* outnumbers *says I* in contemporary data is related to the fact that the tendency towards inversion of pronominal subjects with *say* in reporting clauses, which was so widespread in EModE and relatively well attested in LModE, is no longer visible in PDE. The predominance of the SV pattern in reporting clauses with subject pronouns is a recent development. Figure 5 shows the changes in the frequency of both patterns in the *Corpus of Historical American English* (COHA).<sup>17</sup> It turns out that non-inversion starts to dominate in the latter half of the nineteenth century, while the frequency of the VS pattern with subject pronouns becomes low in the 1930s.

<sup>15</sup>Queries: '*says I* (quotation mark followed by *says I*) and , *says I* (comma followed by *says I*).

<sup>16</sup>Naturally, there are regional differences in the use of non-standard present tense verb endings in English (Beal 2010: 32), and the use of *-s* in the first person singular may be more or less stigmatised depending on various social variables such as education or social class. The instances of *I says* or *says I* in written data come mostly from dialogues in fiction. However, if we assume that the authors try to imitate authentic reporting strategies used by native speakers of English in informal circumstances, the presence of such patterns should not be treated as irrelevant, especially since both patterns are used in corpora of spoken English, see (62).

<sup>17</sup>Since COHA is not syntactically annotated (there are only part-of-speech tags), the search was limited to *said* following a quotation mark and followed by a personal pronoun. PRON-*said* clauses were extracted by means of the query: '*\_pp*\* *said* and *said*-PRON clauses with: '*said* *\_pp*\*. Only clauses with the most frequent pronouns (*I*, *he* and *she*) were taken into account in the calculations; the chart shows the totals for both patterns.

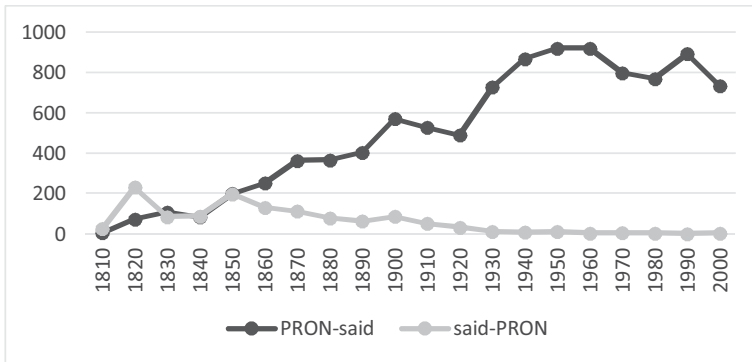


Figure 5. The development of VS and SV parentheticals with *said* and subject pronouns in COHA (normalised frequencies per 1 million words)

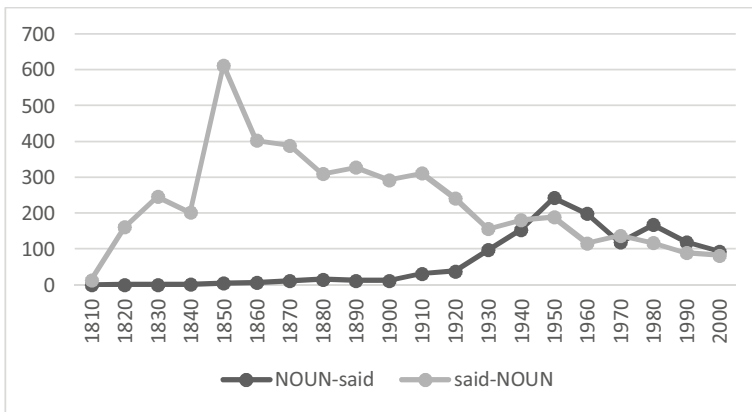


Figure 6. The development of VS and SV parentheticals with *said* and nominal subjects in COHA (normalised frequencies per 1 million words)

This result is especially interesting considering the fact that at around the same time the frequency of non-inversion with nominal subjects<sup>18</sup> started to grow, as shown in figure 6.<sup>19</sup>

<sup>18</sup>For the technical reasons explained in fn. 17, the search was limited to *said* following a quotation mark and followed by a proper noun, the definite article or the indefinite article. NOUN-*said* clauses were extracted by means of the queries: '*said\_np\**' and '*said\_at\**' (only *the*, *a* and *an* were included in the calculation), while *said*-NOUN clauses were extracted by means of: '*\_at\*\_nm\* said*' and '*\_np\* said*' (the chart shows the totals of both patterns).

<sup>19</sup>The data shown in figures 5 and 6 come from American English and therefore some of the differences that we see between the figures from COHA and earlier data from the Penn corpora may be due to dialectal variation. However, since there are no large-scale diachronic corpora of nineteenth- and twentieth-century British English, COHA turned out to be the only available source of such data, which are used here to illustrate the general qualitative tendencies.



It is certainly puzzling why the robust VS pattern started to lose ground in the late nineteenth century, first with pronominal, and then with nominal subjects. A possible mechanism is related to the growing frequency of new reporting verbs and information structure. The pragmatic motivation for the choice between VS and SV in reporting clauses is quite simple: what comes second is more prominent (Biber *et al.* 1999: 922). Placing focus on the verb made little sense if the reporting verb gave no additional information as to the manner or intention of speaking. Therefore, nominal subjects always had a higher information value than *say* or *quoth*, and this may be the pragmatic reason why quotative inversion with nominal subjects was so regularly observed in earlier English. When the subject was pronominal, its information value was as low as that of the reporting verb, so the variation between VS and SV could not depend on information structure (though the choice of the pattern depended on function: reporting preferred VS while commenting was associated with SV). However, in LModE, when new reporting verbs appeared, their clause-final placement could be caused by information structure since the meaning of the verb was often more fine-grained and related to the manner of speaking (*whisper*, *murmur*, *exclaim*, etc.). Thus, their information value was higher than that of pronominal and discourse-old nominal subjects. The results presented in table 10 show that new low-frequency reporting verbs rarely show inversion with pronominal subjects in LModE, which would confirm this tendency. Quirk *et al.* (1985: 1022) report that quotative inversion in PDE is most frequent when the verb is *said*. Since *say* is the simplest reporting verb, its information value is relatively low and the subject is given more prominence by the use of the VS pattern. What is more, non-inversion is reported to be more frequent in fiction (Biber *et al.* 1999: 923), where style and authors' creativity may lead to an increased variety of reporting verbs and strengthen the tendency towards the use of SV. A full analysis of the influence of information structure on the rates of quotative inversion is beyond the scope of this study, but it seems quite plausible that the emergence of new reporting verbs did have an impact on the growing proportion of SV quotative structures in LModE and PDE, with *say* partly joining this general SV pattern in the nineteenth century, by analogy to other verbs with higher information value, and with the *-s* ending generalised to all grammatical contexts (though with the morphological reduction blocked in formal registers by prescriptive pressure).

## 6 The development of the parenthetical reporting construction

According to Traugott & Trousdale (2013), new constructions, defined as pairings of form and meaning organised in a network (Goldberg 2006), emerge as a result of constructionalisation. With time, existing constructions undergo modifications pertaining to their meaning, function, collocational constraints, syntax or morphophonology. In the usage-based approach, a developing construction is said to undergo an increase in productivity and schematicity, and a decrease in compositionality (Traugott & Trousdale 2013: 113–22). Productivity is understood as both token-frequency (the overall number of forms sanctioned by the construction) and

type-frequency (the collocational range of the construction). Schematicity is related to the ‘extensibility’ of the schema (the schema being the highest level of abstraction of a given set of constructions, defined primarily in terms of function), which results in the growing number of micro-constructions integrated into the general schema (Barðdal 2008: 31, quoted in Traugott & Trousdale 2013: 116). A decrease in compositionality is related to the diminishing transparency (and analysability) of the match between the meaning of the construction parts and its form (or syntax). This section aims to account for the development of the reporting parenthetical construction in English within the diachronic construction grammar framework.

The data presented in sections 3–5 show that the parenthetical reporting construction in the form of a quoted message followed by a reporting clause existed already in OE. At first it was neither token-productive (only 130 occurrences per 1 million words) nor type-productive (only two reporting verbs, *cweþan* and *secgan*), and not fully conventionalised because its distribution among various OE texts is uneven. There was also only one verb, *cweþan*, which was attracted to the construction, but the string was fully analysable, without any signs of morphological reduction. However, the fact that *cweþan* was used virtually only in the past tense and with singular subjects led to the predominance of the form *cwæþ*, creating a favourable context for such a reduction, which took place in ME. ME is also the period in which the parenthetical construction with *quoþ* started to appear with the fixed VS order, while in OE there was variation between *he cwæþ* and *cwæþ he* patterns. The fact that this variation was clearly text-based suggests that OE speakers could have made different decisions as to the analysis of ‘quotation + reporting verb’ strings. Some speakers may have analysed such sequences as a part of the V-2 construction, which in OE involved inversion of nominal subjects and non-inversion of subject pronouns (Fischer *et al.* 2000; Ringe & Taylor 2014), so clauses such as (64) and (65), as well as all the parenthetical reporting clauses with *secgan*, are consistent with it.

- (64) Gehyr      ðu,          mann,    **cwæð**    **se**    **halga**    **Ysodorus**,  
 hear-IMP you-SG man    said    the    holy    Isidore  
 ‘Hear, man, said the holy Isidore’ (coverhom,HomU\_7\_[ScraggVerc\_22]:157.2956)
- (65) Ic    gelyfe,    **he**    **cwæð**,    þæt    heo    libbe    þurh    þe.  
 I    believe    he    said    that    she    lives    through    you  
 ‘I believe, he said, that she may live thanks to you’ (coalive,ÆLS\_[Martin]: 494.6275)

However, reporting clauses with inverted pronominal subjects such as (66) do not pattern with OE V-2 clauses because according to the OE V-2 rule only a closed group of short adverbs, especially *þa* ‘then’ and *þonne* ‘then’, as well as the negative particle *ne*, could cause inversion of subject pronouns (Pintzuk 1999: 91; Cichosz 2017a).

- (66) Nese,    **cwæð**    **ic**,    ne    wat    ic    heo  
 no    said    I    not    know    I    her  
 ‘No, I said, I do not know her’ (cobede,Bede\_5:13.430.29.4339)

It is difficult to argue that the quoted message could form a coherent group with these lexical elements (treated as ‘operators’ in formal accounts of OE syntax), especially

given that the inversion mechanism is not consistently applied in all the texts. What is more, Bech (2001) shows that inversion after ‘non-operators’ is more common when nominal subjects are new; given subjects usually stay non-inverted, so the fact that 100 per cent of nominal subjects in OE reporting clauses are inverted is also quite unexpected from the point of view of the OE V-2 because the speakers referred to in reporting clauses are rarely discourse-new.

Another possible interpretation of (64) and (66) would be to analyse them as instances of V-1 (as in Calle-Martín & Miranda-García 2010). However, this interpretation is not unproblematic because V-1 main declarative clauses are used in OE for a number of pragmatic reasons: such clauses mark transition to a new action, summarise the discussion, introduce a new/contrasting character/type and open a new story/paragraph (Ohkado 2004). Reporting parenthetical clauses fulfil none of these functions so their treatment as representatives of V-1 declaratives is unconvincing; from a pragmatic point of view, V-1 was a different construction.<sup>20</sup>

In sum, (64) and (65) may be analysed as instances of V-2, whereas (64) and (66) are sometimes analysed as V-1, but neither of these interpretations explains the existence of all three closely related and functionally equivalent patterns shown in (64)–(66). However, we need to remember that (65) and (66) do not co-occur in the same texts. This suggests that while for some OE speakers (64) and (65) could be linked to the V-2 construction, for other speakers (e.g. the Bede translator), strings such as (64) and (66) were part of a different construction (‘quotation + *cweþan* + subject’), which had a different function than the OE V-2 (or V-1), was not a part of the V-2 schema and thus followed its own syntax. Therefore, we could explain the variation between SV and VS observed in OE by the ongoing constructionalisation process of the VS parenthetical reporting construction, which was reanalysed by some speakers as having no subject-type constraints, while others still linked it with the V-2 construction in which the difference in the behaviour of pronominal and nominal subjects was observed.

In ME, the token-frequency of the parenthetical reporting construction undergoes a considerable increase. The central lexeme co-occurring with the construction is still *quethen*, whose association with the construction is strengthened and the verb undergoes morphological reduction, thus making *quod/quop* less analysable and leading to the decreased compositionality of the string *quod/quop* + subject. Moreover, *seien* outnumbers *quethen* in the construction, though as a generally frequent verb it does not show a clear preference for parenthetical use. In this period, the difference

<sup>20</sup> V-1 declarative clauses are relatively rare in OE prose (Allen 1995: 34; Calle-Martín & Miranda-García 2010; Mitchell 1985: §3930). In Cichosz (2017c) it is reported that there are only c. 900 V-1 clauses in the YCOE corpus, which means that including or excluding clauses with quotative inversion such as (64) or (66) may have a significant influence on the analysis. This study identified 144 clause-medial and clause-final reporting parentheticals with the VS order; their inclusion in the study of the OE V-1 would increase the number of such clauses by 14 per cent. What is more, if clauses with quotative inversion are included, the association between the presence of verbs of saying and the use of the V-1 order is strengthened. In Cichosz (2017b) it is shown that this association is rather weak in all OE prose texts except Bede if clauses with quotative inversion are excluded.

in the behaviour of nominal and pronominal subjects is non-existent with *quethen* and gradually disappearing with *seien*, which shows that more and more speakers were linking both reporting verbs to the VS parenthetical reporting construction. The early ME cases of SV with *seien* may be seen as a continuation of the pattern identified in OE for *cweþan*, i.e. the still existing link between reporting parenthetical clauses and the V-2 construction.

In EModE, the collocability of the construction rises, with new reporting verbs joining the pattern. What is more, *say* becomes attracted to the construction, though *quoth* still shows the strongest association with it. The tendency towards inversion is increased and generalised to all subject types, leading to a significant increase in inversion rates with subject pronouns, most probably resulting from a reorganisation and split of the parenthetical construction with verbs of saying into the reporting construction (with quotative inversion) and the comment construction (following the SV pattern), as shown in section 6. In this period, inversion was a clear sign of the function of parenthetical clauses, with no link between quotative inversion and V-2 (which was ‘largely lost during the late Middle English period’ (Fischer *et al.* 2000: 105)).

In LModE, the type-frequency of the construction is considerably increased. The most attracted lexical item is *say* (*quoth* is marginalised and becomes obsolete), but there are more new verbs associated with it, so we may conclude that the construction keeps extending its collocational range. However, the EModE generalisation of quotative inversion as a marker of reporting function is no longer observed; both the VS and the SV pattern exist within the parenthetical reporting construction (the schema is extended). The speakers reanalyse SV parentheticals with *say* as reporting especially when the invariant *-s* ending is used for present reference (and *-ed* for past reference, but in this context no functional overlap with comment clauses was possible), while comment clauses show the standard endings (*I say*). This reanalysis is not fully reflected in written data because in LModE, when English grammar was already standardised, the prescriptive pressure to use standard endings was too strong for full morphological reduction to take place (unlike in ME, when *quethen* could undergo such a process in relatively unrestrained conditions).

## 7 Conclusion

The analysis presented in this article shows that the parenthetical reporting construction has been used in English since the OE period, and it has followed its own constructionalisation path, with various constructional changes including growing productivity and collocability as well as morphological reduction of *quoth* and (partly) *say*. The investigation also shows that clauses with quotative inversion, regardless of some superficial similarity and a partial structural overlap, pattern neither with clauses following the V-2 rule nor with V-1 declaratives. This confirms the assumptions of Los (2009) and Haerberli (2002) that such clauses should not be included in studies of the V-2 rule. What is more, the analysis presented above suggests that the parenthetical reporting construction did not emerge from the V-2 construction. In OE and ME both

constructions existed alongside each other (and interacted with each other, as shown in section 6), while in EModE the V-2 rule was lost while quotative inversion became generalised as the main pattern for the parenthetical reporting clauses with no subject type constraints. The fact that PDE shows variation between SV and VS is a recent development, and the analysis suggests that it may have been triggered by the increase in the number of reporting verbs used in the construction and their high information value, though a comprehensive study of nineteenth- and twentieth-century English is needed to confirm this hypothesis.

*Author's address:*

*Department of English and Applied Linguistics  
Institute of English Studies  
University of Łódź  
ul. Pomorska 171/173 90–236 Łódź  
Poland  
[an.cichosz@gmail.com](mailto:an.cichosz@gmail.com)*

#### References

- Allen, Cynthia L. 1995. *Case marking and reanalysis*. Oxford: Oxford University Press.
- Ambar, Manuela. 1992. *Para uma sintaxe da inversão sujeito-verbo em português*. Lisbon: Colibri.
- Bailyn, John Frederick. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press.
- Barbieri, Federica. 2005. Quotative use in American English: A corpus-based, cross-register comparison. *Journal of English Linguistics* 33, 222–56.
- Barðdal, Johanna. 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.
- Beal, Joan C. 2010. *An introduction to regional Englishes*. Edinburgh: Edinburgh University Press.
- Bech, Kristin. 2001. Word order patterns in Old and Middle English: A syntactic and pragmatic study. PhD dissertation, University of Bergen.
- Bech, Kristin. 2017. Old truths, new corpora: Revisiting the word order of conjunct clauses in Old English. *English Language and Linguistics* 21(1), 1–25.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Bonami, Olivier & Danièle Godard. 2008. On the syntax of direct quotation in French. In Stefan Müller (ed.), *Proceedings of the HPSG08 Conference*, 355–77. Stanford, CA: CSLI Publications.
- Brinton, Laurel J. 2008. *The comment clause in English: Syntactic origins and pragmatic developments*. Cambridge: Cambridge University Press.
- Bruening, Benjamin. 2016. Alignment in syntax: Quotative inversion in English. *Syntax* 19(2), 111–55.
- Buchstaller, Isabelle. 2013. *Quotatives: New trends and sociolinguistic implications*. Oxford: Wiley-Blackwell.
- Calle-Martín, Javier & Antonio Miranda-García. 2010. ‘Gehyrdon ge þæt gecweden wæs’: A corpus-based approach to verb-initial constructions in Old English. *Studia Neophilologica* 82, 49–57.

- Cheshire, Jenny & Sue Fox. 2008. Performed narrative: The pragmatic function of 'this is me' and other quotatives in London adolescent speech. Paper presented at the 17th Sociolinguistics Symposium (3–5 April 2008, Amsterdam).
- Cichosz, Anna. 2017a. Inversion after clause-initial adverbs in Old English: The special status of *þa*, *þonne*, *nu* and *swa*. *Journal of English Linguistics* 45(4), 308–37.
- Cichosz, Anna. 2017b. Old English verbs of saying and verb-initial order. *English Language and Linguistics* 21(3), 533–58.
- Cichosz, Anna. 2017c. Verb-initial main clauses in Bede: A translation effect? In Andrzej Kijak, Jerzy Nykiel & Andrzej M. Łęcki (eds.), *Current developments in English historical linguistics*, 68–89. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. *Audio BNC: The audio edition of the Spoken British National Corpus*. Phonetics Laboratory, University of Oxford. [www.phon.ox.ac.uk/AudioBNC](http://www.phon.ox.ac.uk/AudioBNC)
- Collins, Chris & Phil Branigan. 1997. Quotative inversion. *Natural Language and Linguistic Theory* 15, 1–41.
- Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 520 million words, 1990–present*. <http://corpus.byu.edu/coca/>
- Davies, Mark. 2010–. *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. <http://corpus.byu.edu/coha/>
- Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. *The syntax of early English*. Cambridge: Cambridge University Press.
- Gärtner, Hans-Martin & Beáta Gyuris. 2014. A note on quotative inversion in Hungarian. *Finno-Ugric Languages and Linguistics* 3, 2–30.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Gries, Stefan Th. 2007. *Coll.analysis 3.2a*. A program for R for Windows 2.x.
- Griffiths, James. 2015. Speaker and quote reduced parenthetical clauses. In Stefan Schneider, Julie Glikman & Mathieu Azanzi (eds.), *Parenthetical verbs*, 71–102. Berlin: Mouton De Gruyter.
- Haerberli, Eric. 2002. Observations on the loss of Verb Second in the history of English. In C. Jan-Wouter Zwart & Werner Abraham (eds.), *Studies in comparative Germanic syntax*, 245–72. Amsterdam: John Benjamins.
- Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann & Ian H. Witten. 2009. The WEKA data mining software: An update. *SIGKDD Explorations* 11(1).
- Harbert, Wayne. 2007. *The Germanic languages*. Cambridge: Cambridge University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Joseph, Brian D. 2000. Textual authenticity: Evidence from medieval Greek. In Susan C. Herring, Pieter van Reenen & Lene Schøsler (eds.), *Textual parameters in older languages*, 309–30. Amsterdam: John Benjamins.
- Kemenade, Ans van. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.
- Konno, Masatoshi. 2013. Inverted reporting verbs in the sentence-initial position in English. *Token: A Journal of English Linguistics* 2, 149–68.
- Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. *The Penn–Helsinki Parsed Corpus of Early Modern English (PPCEModE)*. Department of Linguistics, University of Pennsylvania.
- Kroch, Anthony, Beatrice Santorini & Ariel Diertani. 2016. *The Penn Parsed Corpus of Modern British English (PPCMBE2)*. Department of Linguistics, University of Pennsylvania.

- Kroch, Anthony & Ann Taylor. 2000. *The Penn–Helsinki Parsed Corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania.
- Lampert, Martina. 2013. *Say, be like, quote (unquote)*, and the air-quotes: Interactive quotatives and their multimodal implications. *English Today* 29(4), 45–56.
- Los, Bettelou. 2009. The consequences of the loss of verb-second in English: Information structure and syntax in interaction. *English Language and Linguistics* 13(1), 97–125.
- Matos, Gabriela. 2013. Quotative inversion in Peninsular Portuguese and Spanish, and in English. *Catalan Journal of Linguistics* 12, 111–30.
- Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Clarendon Press.
- Moore, Colette. 2011. *Quoting speech in early English*. Cambridge: Cambridge University Press.
- Moore, Colette. 2015. Histories of talking about talk: *quethen, quoth, quote*. In Jenny Aendholz, Wolfram Bublitz & Monika Kirner-Lundwig (eds.), *The pragmatics of quoting now and then*. Berlin: Mouton de Gruyter.
- Ohkado, Masayuki. 2004. On the structure and function of V1 constructions in Old English. *English Studies* 85, 2–16.
- Pezik, Piotr. 2015. Spokes – a Search and Exploration Service for Conversational Corpus Data. In *Selected Papers from the CLARIN 2014 Conference*, 24–25 October 2014, Soesterberg, the Netherlands, 99–109. Linköping Electronic Conference Proceedings. Linköping University Electronic Press, Linköpings Universitet.
- Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland.
- Quinlan, John Ross. 1993. *C4. 5: Programs for machine learning*. San Francisco: Morgan Kaufmann.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- R Core Team. 2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. [www.R-project.org](http://www.R-project.org)
- Randall, Beth, Anthony Kroch & Ann Taylor. 2005–13. *CorpusSearch 2*.
- Ringe, Don & Ann Taylor. 2014. *A linguistic history of English*, vol. II: *The development of Old English*. Oxford: Oxford University Press.
- Rohrbacher, Bernhard. 1994. English main verbs move never. *Penn Review of Linguistics* 18, 145–59.
- Sams, Jessie. 2009. Genre-controlled constructions in written language quotatives: A case study of English quotatives from two major genres. In Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen Wheatley (eds.), *Formulaic language 1: Distribution and historical change*, 147–69. Amsterdam: John Benjamins.
- Sonoda, Kenji. 1997. Subject–Verb inversion before a quotation in the media discourse. *Current English Studies* 36, 13–25.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2), 209–43.
- Suñer, Margarita. 2000. The syntax of direct quotes with special reference to Spanish and English. *Natural Language and Linguistic Theory* 18, 525–78.
- Tagliamonte, Sali & Alex D’Arcy. 2004. *He’s like, she’s like*: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8(4), 493–514.
- Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *The York–Toronto–Helsinki Parsed Corpus of Old English Prose (YCOE)*. Department of Linguistics, University of York. Oxford Text Archive.
- Traugott, Elizabeth Closs. 2012. Middle English: Pragmatics and discourse. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics: An international handbook*, 466–79. Berlin: De Gruyter Mouton.

- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Westergaard, Marit. 2007. English as a mixed V2 grammar: Synchronic word order inconsistencies from the perspective of first language acquisition. *Poznań Studies in Contemporary Linguistics* 43(2). 107–31.
- Zwart, C. Jan-Wouter. 2005. Verb second as a function of merge. In Marcel den Dikken & Christina M. Tortora (eds.), *The function of function words and functional categories*, 11–40. Amsterdam: John Benjamins.