DISTRIBUTIONAL DIFFERENCES BETWEEN OLD ENGLISH MAIN CLAUSES WITH AND WITHOUT A CONJUNCTION

Richard Zimmermann
University of Geneva

Proceedings of the LFG14 Conference

Miriam Butt and Tracy Holloway King (Editors)

2014

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

Old English main clauses and main clauses introduced by a conjunction pattern differently in terms of verb placement and topicalization. An LFG model is proposed to capture these differences. It allows conjunctions to be inserted under C, thereby blocking the CP-layer as a locus for finite verbs and topics. The model is supported by a statistical examination of relevant word order patterns in the extant Old English text corpus.

1 Introduction

A distinction is commonly made between Old English ordinary main clauses and those main clauses that are introduced by a conjunction like *and* 'and', *ac* 'but' or *ne* 'nor' (e.g., Mitchell 1985: 694, 967; Bech 2001: 86-93). The reason is that there are two pervasive word order differences between these two clause types.

Firstly, in contrast to main clauses (henceforth MC), main clauses with a conjunction (henceforth CC) show substantial rates of verb-final patterns, as in (1).[1]

(1) **&** hie þa Romane be þara biscepa lare hie swa cuce *bebyrgdan*
and they the Romans by the bishops' lore them thus alive buried
'And the Romans buried them alive according to the bishops' teaching'
(coorosiu,Or_4:7.98.9.2019)

Secondly, preposing of a non-subject constituent in front of a subject, commonly referred to as 'topicalization', is more common in MCs than in CCs. That is to say, constituent fronting is more likely in a sentence such as (2a), where the topic appears in clause-initial position, than in a sentence like (2b), which is introduced by a conjunction.

(2) a. [$_{DP}$ Todældu wæteru] we lætað ut of urum eagum
divided waters we let out of our eyes
'We let the divided waters out of our eyes'
(cocura,CP:53.413.27.2858)

b. **and** [$_{DP}$ þone sang] we sungon unseldon mid heom
and the song we sang frequently with them
'And we frequently sang the song with them'
(coaelive,ÆLS_[Swithun]:262.4384)

†I am grateful for helpful comments from the audience of the LFG 2014 conference in Ann Arbor, MI, Cynthia Allen for a very useful review, an anonymous reviewer, Steffen Eisner, the UPenn Linguistics Department's Treebanks Group - in particular Aaron Ecay - and my supervisor, Prof. Eric Haeberli. Thanks are also due to Prof. Paul R. Rosenbaum for his suggestion to use odds ratios as a measure of effect size. All remaining errors are my own.

[1]Many Old English manuscripts include the so-called 'Tironian note *ond*' for the word 'and.' It looks like '7' but is rendered here as the ampersand sign '&.'

In this paper, I will present an LFG model of Old English main clauses with and without conjunctions that accounts for these distributional differences. Section 2 formalizes the verb placement facts. Section 3 supports the proposed model with empirical data from the extant Old English text corpus. Sections 4 and 5 do the same for the topicalization facts. Section 6 concludes.[2]

2 Formal Analysis I: Verb Placement

In this section, I formalize the distributional difference in verb placement between Old English MCs and CCs. I will first analyze possible verb positions in Old English, then introduce the concept of C-head conjunctions and finally show how the resulting model accounts for the observed clause type effect.

2.1 Verb Positions in Old English

I will follow the analysis of Old English verb placement patterns as laid out in Pintzuk (1999). Subject pronouns play an important role in this description as diagnostic elements because they can occur in only one fixed position, namely the specifier of `IP`. Pronominal subjects are underlined in the following examples.

The finite verb in "ordinary" main clauses, i.e. in positive, indicative declaratives, is positioned in `I`, the lowest possible position for finite verbs in Old English. Pronominal subjects will therefore unvaryingly occur before the finite verb and after preposed constituents, (3a). Full subjects, on the other hand, can be placed in a lower position, the specifier of `VP`, creating verb-second patterns, (3b).

(3) a. þis bebod [$_{IP}$ ic *nam* æt minum Fæder]
this command I took at my father
'I received this command from my father'
(cowsgosp,Jn_[WSCp]:10.18.6634)

b. ðæne aþ [$_{IP}$ *nam* Wulfsige se scirigman]
the oath took Wulfsige the sheriff
'Sheriff Wulfsige received the oath'
(codocu3,Ch_1458_[Rob_41]:36.50)

The head `I` can be variably headed with respect to its complement, `VP`, i.e. the model of Old English can generate both head-final (e.g. object-verb) as well as head-initial (e.g. verb-object) `IP`s, (4). This is true for main as well as subordinate clauses. For example, (5a) illustrates a head-initial, (5b) a head-final `IP`. The particle *ut* 'out' is a diagnostic identifying the position of `VP`.

[2]The syntactically parsed *York Corpus of Old English, YCOE* (Taylor et al. 2003) is the source for all examples, their citation, and the statistical counts that the generalizations described here are based on. I followed the annotation scheme of the YCOE to determine whether a clause was to be counted as main, coordinated, or subordinate. Appendix 1 lists all electronic text files that were included in the quantitative analyses.

(4) $\texttt{I}' \rightarrow \{ \begin{array}{ll} \texttt{VP} & \texttt{I} \\ \uparrow=\downarrow & \uparrow=\downarrow \\ \mid\ \texttt{I} & \texttt{VP} \\ \uparrow=\downarrow & \uparrow=\downarrow \} \end{array}$

(5) a. þa yflan [$_{\text{IP}}$ hig *awurpon* [$_{\text{VP}}$ ut]]
the evil they threw out
'They threw out the bad ones' (cowsgosp,Mt_[WSCp]:13.48.900)

b. manega deofolseocnyssa [$_{\text{IP}}$ he [$_{\text{VP}}$ ut] *adraf*]
many devil-sicknesses he out drove
'He drove out many demoniacal possessions'
(cowsgosp,Mk_[WSCp]:1.34.2248)

In contrast, a few contexts allow the finite verb to occur high in C, thus preceding pronominal subjects. These contexts are clauses with interrogative force, (6a), or with initial presentational adverbs like *þa* and *þonne* 'then,' (6b), where high verb placement is essentially categorical. Furthermore, negative verbs, (6c), and subjunctive verbs tend to be placed high, (6d), but may also occur in I.[3]

(6) a. hwanon *come* þu Giezi?
whence come you Gehazi
'Where do you come from, Gehazi?'
(cocathom1,ÆCHom_I,_27:408.241.5443)

b. þa *cwoman* we to sumre byrig
then came we to some city
'Then we arrived in a city' (coalex,Alex:15.1.126)

c. *Ne æt* he næfre flæsc
not ate he never flesh
'He didn't ever eat meat'
(comart3,Mart_5_[Kotzor]:Ju22,A.8.1012)

d. *Lufian* we urne Sceppend
love.subjunctive we our creator
'Let us love our creator'
(coblick,HomU_18_[BlHom_1]:5.51.50)

[3]Other contexts may allow high verb placement as well. For example, high verbs may occur in imperatives, verb-first conditionals or particularly dramatic and lively narratives (e.g., Kemenade 1987: 44-5) as in (i). However, these contexts are relatively rare. I will therefore ignore them here.

(i) *Comon* hi of þrim folcum, ðam strangestan Germanie
came they from three peoples the strongest of-Germany
'They came from the three strongest peoples of Germany'
(cobede,Bede_1:12.52.2.469)

The following annotated phrase structure rules captures these generalizations.

(7) a. C → NEG C
↓ ∈ (↑ ADJ) ↑ = ↓

b. C → V
↑ = ↓
{ (↑ CLAUSE-TYPE)= int
| (↑ UDF ADJ-TYPE)=$_c$ pres-top
| (↑ ADJ ∈ ADJ-TYPE)=$_c$ neg
| (↑ MOOD)=$_c$ subjunctive }

If a finite verb is placed under C, the clause is interpreted as interrogative. Alternatively, finite verbs are licensed under C if there exists a clause-initial underspecified discourse function, UDF, that contains an adjunct type pres-top feature introduced by adverbs like *þa* or *þonne*. Mutually exclusive clause-type and UDF adjunct type requirements on I will rule out low verb placement in these contexts thus ensuring the obligatory nature of these constraints. Finally, a verb may appear in C if a member of the set of adjuncts includes a neg feature or if the verb's mood is subjunctive. Since no parallel negative constraints are found on I, negative or subjunctive verbs may also be placed low.

2.2 Dual Category Membership of Conjunctions

I propose that Old English conjunctions belong to two distinct categories. They may function as *logical connectors* as in Modern English. In this case, they are introduced as category CONJ by a phrase structure rule for sentential coordination, which builds a coordinate structure with a set of individual conjunct clauses, (8a). The conjunction bears a non-distributive COORD-FORM feature, (8b).

(8) a. CP → CP$^+$ CONJ CP
↓ ∈ ↑ ↑ = ↓ ↓ ∈ ↑

b. *and* CONJ (↑ COORD-FORM)= and

Alternatively, Old English conjunctions can be of the same type as complementizers, i.e. C, and I therefore refer to those items as *C-head conjunctions*. A set of X-bar compliant phrase structure rules constructs the Old English clausal left periphery, headed by C, (9a). Since conjunctions are of the same category, (9b), they may be inserted under this position.

(9) a. CP → C′ C′ → C IP
↑ = ↓ ↑ = ↓ ↑ = ↓

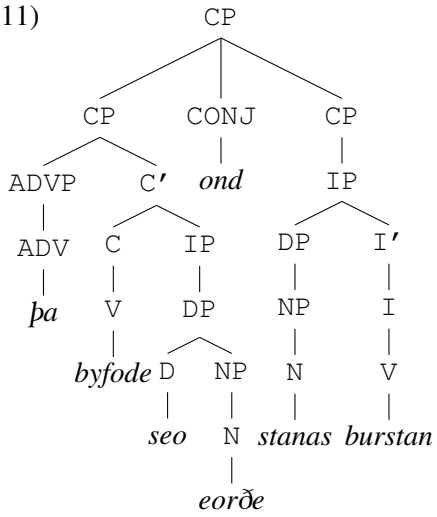
b. *and* C (↑ COORD-FORM)= and

The final CP-clausal conjunct in (8a) can be annotated with a negative existential constraint, $\neg(\downarrow$ COORD-FORM$)$. This rules out simultaneous application of both logical connectors and C-head conjunctions in the same structure.

Old English constructions of the form 'clause - conjunction - clause' will now be ambiguous between parses of the conjunction as a logical connector or as a C-head conjunction. For example, the sentence in (10) can be analyzed as a coordinate structure with a logical connector, as in (11), or as two independent clauses each with their own f-structure, as shown in (12).

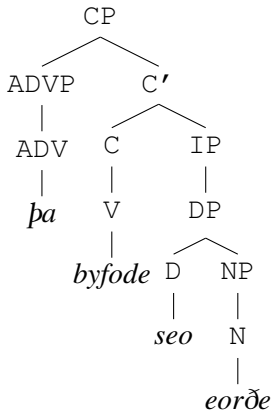
(10) þa byfode seo eorðe ond stanas burstan
then quaked the earth and stones burst
'Then the earth quaked and stones burst'
(comart3,Mart_5_[Kotzor]:Ma25,A.15.459-460)

(11)

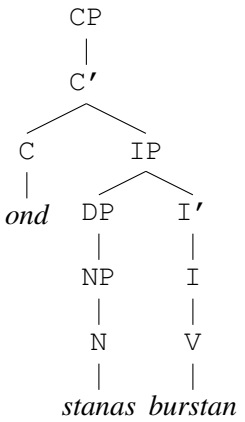


$$\left[\left\{\left[\begin{array}{ll}\text{PRED} & \text{‘quake⟨SUBJ⟩’}\\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘earth’}\\ \text{DEF} & +\\ \text{CASE} & \text{nom}\\ \text{NUM} & \text{sg}\\ \text{GENDER} & \text{fem}\end{array}\right]\\ \text{ADJ} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{‘then’}\\ \text{ADJ-TYPE} & \text{pres-top}\end{array}\right]\right\}\\ \text{TENSE} & \text{past}\\ \text{MOOD} & \text{ind}\end{array}\right],\left[\begin{array}{ll}\text{PRED} & \text{‘burst⟨SUBJ⟩’}\\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘stone’}\\ \text{CASE} & \text{nom}\\ \text{NUM} & \text{pl}\\ \text{GENDER} & \text{masc}\end{array}\right]\\ \text{TENSE} & \text{past}\\ \text{MOOD} & \text{ind}\end{array}\right]\right\},\ \text{COORD-FORM}\ \text{and}\right]$$

(12)



$$\left[\begin{array}{ll}\text{PRED} & \text{‘quake⟨SUBJ⟩’}\\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘earth’}\\ \text{DEF} & +\\ \text{CASE} & \text{nom}\\ \text{NUM} & \text{sg}\\ \text{GENDER} & \text{fem}\end{array}\right]\\ \text{ADJ} & \left\{\left[\begin{array}{ll}\text{PRED} & \text{‘then’}\\ \text{ADJ-TYPE} & \text{pres-top}\end{array}\right]\right\}\\ \text{TENSE} & \text{past}\\ \text{MOOD} & \text{ind}\end{array}\right]$$

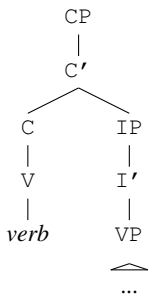


$$\left[\begin{array}{ll}\text{PRED} & \text{‘burst⟨SUBJ⟩’}\\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘stone’}\\ \text{CASE} & \text{nom}\\ \text{NUM} & \text{pl}\\ \text{GENDER} & \text{masc}\end{array}\right]\\ \text{TENSE} & \text{past}\\ \text{MOOD} & \text{ind}\\ \text{COORD-FORM} & \text{and}\end{array}\right]$$

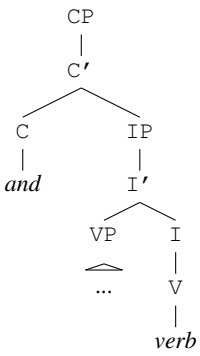
2.3 Accounting for the Distributional Difference in Verb Placement

The proposed analysis accounts for the fact that CCs are more commonly verb-final than MCs as follows. C is a potential verb position in all MCs (in the contexts outlined above), (13a). C-head conjunctions, however, may block the verb's appearance under C. Consequently, C-head conjunctions will force higher rates of structures with the finite verb in I than MCs in general, and higher rates of verb-final patterns in particular, (13b). C-head conjunctions in CCs behave exactly like complementizers in subordinate clauses in this respect. However, the complementarity between high verbs in C (in the contexts outlined above) and the presence of conjunctions is merely a tendency and not categorical (unlike the complementarity between verbs in C and complementizers in subordinate clauses). The reason is that conjunctions are not necessarily inserted under C but may be logical connectors instead and can then co-occur with verbs in C, (13c). Hence, CCs will show more verb-final structures than MCs but fewer than subordinate clauses.[4]

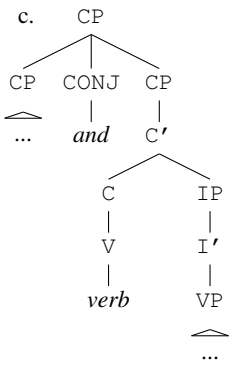
(13) a.



b.



c.



3 Statistical Evidence I

3.1 Source of Surplus of Verb-Final Structures in CCs

The above model requires that CCs should have a higher rate of I-final word order specifically at the expense of constructions with high verb placement and not because of an inherently higher rate of I-final headedness. I assessed the validity of this assumption by counting three types of MCs and CCs in the surviving Old English texts. The first type shows the finite verb immediately preceding a subject pronoun, (14). This *VS...* pattern indicates high verb placement.

[4]The distributional difference in verb placement between MCs and CCs follows largely from the same structural configurations as in den Besten's (1983) classical analysis of the verb-second constraint in Modern West Germanic. The actual rates of I-initial and I-final headedness in MCs and CCs can be assumed to be basically identical.

(14) a. *Ne geherst* [$_{IP}$ þu þas word fram me]
not hear you these words from me
'You do not hear these words from me'
(coalcuin,Alc_[Warn_35]:435.331)

b. **&** þa *æteawde* [$_{IP}$ he hine Nerone þy þriddan dæge]
and then appeared he reflexive Nero the third day
'And then he appeared to Nero on the third day'
(coblick,LS_32_[PeterandPaul[BlHom_15]]:183.221.2364)

The second type includes any kind of overt subject, a subsequent finite verb, where the two elements are not immediately adjacent. Instead, there is an intervening VP-constituent such as a non-pronominal object, non-finite verb, predicate, particle etc., as in (15). This *S...V* pattern is indicative of I-final headedness.

(15) a. Moyses [$_{I'}$ ða Godes hæse [$_{I}$ *gefylde*]]
Moses then God's order fulfilled
'Then Moses fulfilled God's order'
(cocathom2,ÆCHom_II,_12.1:114.133.2471)

b. **and** his geferan [$_{I'}$ ða mid fleame [$_{I}$ *ætburston*]]
and his companions then with flight escaped
'And his companions then escaped by flight'
(cocathom2,ÆACHom_II,_14.1:141.114.3124)

The third type consists of any kind of overt subject with an immediately following finite verb, (16). Negation, light adverbs or pronouns may intervene between subject and verb. An example with an intervening pronoun is shown in (17). This *SV* pattern is compatible with an I-initial parse.

(16) a. þæt word [$_{I'}$ [$_{I}$ *is*] ælmihtig God]
that word is almighty God
'That word is the almighty God'
(cocathom1,ÆCHom_I,_25:385.189.4940)

b. **&** þæt word [$_{I'}$ [$_{I}$ *wæs*] God]
and that word was God
(cocathom1,ÆCHom_I,_2:195.166.421)

(17) a. Philippus [him [$_{I}$ *dyde*] heora wig unweorð]]
Phillipus for-himself did their warfare unworthy
'Philip considered their warfare unworthy of his attention'
(coorosiu,Or_3:7.64.28.1264)

b. **ne** nenig man [hine [$_{I}$ *geseah*] swiðe hlahendne]]
nor no man him saw very laughing
'Nor did any man see him laughing very much'
(cochad,LS_3_[Chad]:233.151)

The results of this investigation are presented in Table 1 below. It also includes the counts for subordinate clauses (SC) but only for comparative purposes.

Clause type	High verb *V S ...*	Verb-final *S ... V*	Other *S V*
MC	6,066 (19.5%)	1,822 (5.9%)	23,147 (74.6%)
CC	952 (5.2%)	2,952 (16.2%)	14,362 (78.6%)
SC	37 (0.1%)	9,844 (39.3%)	15,169 (60.6%)

Table 1: Verb Positions in MCs, CCs and SCs

As Table 1 shows, 19.5% of all MCs in the sample place the verb before a subject pronoun and hence under C while only 5.9% of all MCs are verb-final. The pattern is essentially inverted in CCs. Only 5.2% of all CCs are indicative of high verb placement[5] whereas 16.2% of all CCs are I-final. This difference is significant (χ^2=2987.0, df=1, p<0.001, odds ratio = 10.32, 95% confidence intervals: 9.43 – 11.29) and suggests that CCs do indeed manifest more verb-final structures than MCs precisely because of a reduction in high verb placement. In contrast, there is a much smaller difference between these patterns and clauses with (potentially) I-initial phrase structure, comprising 74.6% of all MCs and 78.6% of all CCs respectively. A chi-square test reveals that this difference, too, is statistically significant. However, the clause type effect is weak as measured by the odds ratio. Specifically, the odds of finding a (potentially) I-initial sentence is only about 1.25 times greater for CCs as compared to those for MCs (χ^2=103.1, df=1, p<0.001, odds ratio = 1.25, 95% confidence intervals: 1.20 – 1.31). This finding supports the assumption that the actual rates of the variants of IP-headedness do not substantially differ between MCs and CCs.

3.2 Effect of Separation on I-final Clauses

The formal analysis given above entails a differential realization of IP-headedness between clauses introduced by conjunctions that are separated from a subject pronoun and thus must necessarily be logical connectors, (18a), and those that are immediately adjacent to a pronominal subject and can thus potentially be analyzed as C-head conjunctions, (18b). If the conjunction must be a logical connector, the frequency of I-final structures should not differ substantially between CCs and MCs. If the conjunction can be analyzed as a C-head, more I-final structures are expected in CCs than in MCs.

[5]It is not the case that MCs place the verb more frequently under C than CCs simply because they have more negation overall. In fact, negation is somewhat more probable in CCs (10.2%, 1,831/17,985 examples) than in MCs (7.5% 2,280/30,293 examples). A similar assessment for the effect of verbal mood is hard to obtain since the corpus annotation for subjunctives is relatively inaccurate (Ann Taylor, p.c.). I am grateful to Aaron Ecay for raising this point.

(18) a. *CC-separated: necessarily logical connector*

& [PP on ðam seofoðan dæge] he *geendode* his weorc
and on the seventh day he ended his work

'And on the seventh day, he finished his creation'
(cocathom1,ÆCHom_I,_1:182.95.90)

b. *CC-adjacent: potential C-head conjunction*

and he ða mid geleafan his lif *geendode*
and he then with belief his life ended

'And he then ended his life with faith'
(coaelive,ÆLS_[Maccabees]:104.4880)

I collected all main clauses with pronominal subjects and finite verbs. For each example, I recorded the variable 'clause type' with the variants MC, separated CC and adjacent CC. In separated CCs, some constituent, including finite verbs, intervenes between conjunction and pronominal subject. In adjacent CCs, the conjunction is immediately followed by the subject pronoun. I then observed whether the finite verb appeared in a configuration that indicates I-final headedness or not.

Table 2 below summarizes the results of this investigation. As for the previous study, the table also includes the counts for subordinate clauses (SC) to illustrate the fact that, regarding the rate of I-final headedness, CCs fall in between this clause type and MCs.

Clause type	I-final	Other
MC	717 (4.1%)	16,926 (95.9%)
CC-separated	266 (7.6%)	3,216 (92.4%)
CC-adjacent	1,523 (20.3%)	5,986 (79.7%)
SC	7,289 (41.4%)	10,345 (58.6%)

Table 2: I-final Headedness in MCs, separated CCs, adjacent CCs and SCs

As expected, I-final clauses are significantly more common in clauses with potential C-head conjunctions, 20.3%, than in MCs, 4.1% (χ^2=1763.3.6, df=1, $p<0.001$, odds ratio = 6.00, 95% confidence intervals: 5.47 – 6.60) or in clauses that necessarily include logical connectors, 7.6% (χ^2=278.1, df=1, $p<0.001$, odds ratio = 3.08, 95% confidence intervals: 2.68 - 3.53). In contrast, there is a much smaller, although significant, difference in the rate of I-final headedness between MCs and separated CCs, which has a much smaller effect size as measured by the odds ratio (χ^2=83.0, df=1, $p<0.001$, odds ratio = 1.95, 95% confidence intervals: 1.68 – 2.26). These statistics thus provide further evidence for the proposed analysis, in which the high rate of I-final CCs predominantly result from conjunctions that block C as a verb position and not from the presence of conjunctions generally.

4 Formal Analysis II: Topicalization

This section deals with the formal implementation of topicalization in Old English. As before, I will first outline the formal model and then show how it entails the distributional difference in the rate of topicalization between MCs and CCs.

4.1 Topicalization in Old English

It is not uncommon for Old English heavy constituents to be placed in a discourse function in clause-initial position, a phenomenon commonly referred to as 'topicalization.' I use the term here as a convenient way of referring to any configuration that involves constituent fronting to the clausal left periphery (for discussions of the information structure of Old English topics, see e.g., Speyer 2010; Light 2012).

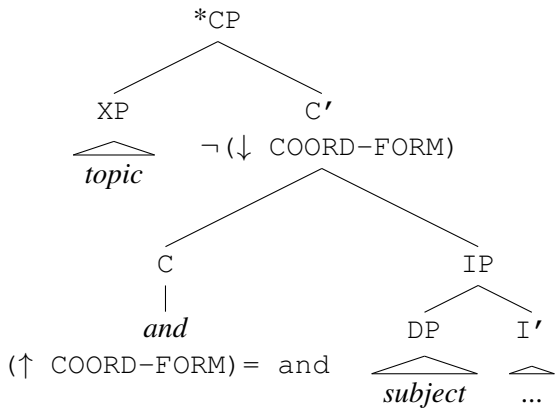
I model topicalization with the annotated phrase structure rule in (19). An underspecified discourse function, UDF, in the specifier of CP must be unified with a grammatical function, which may be arbitrarily deeply embedded within complemented clauses.

(19) $\mathrm{CP} \rightarrow$ XP ($\uparrow$ UDF) = $\downarrow$; C′ $\uparrow = \downarrow$

($\uparrow$ UDF) = ($\uparrow$ $\{$XCOMP | COMP$\}^*$ GF) ; $\neg$($\downarrow$ COORD-FORM)

I constructed the model in such a ways as to rule out parses with a topic in the specifier of CP and, simultaneously, a conjunction under C. This is achieved by annotating C′ with a negative existential constraint, $\neg$($\downarrow$ COORD-FORM), which conflicts with the feature specification of C-head conjunctions. Hence, C-head conjunctions will be incompatible with the topicalization rule and, as desired, structures of the form 'topic - conjunction - subject' cannot be generated, (20).

(20)



It might be objected that this constraint makes the model too restrictive as there are examples attested in Old English that might be compatible with a parse that involves a clause-initial topic followed by a potential C-head conjunction. For example, (21) includes the conjunction, *and*, preceded by a nominative left-dislocation, *John the evangelist*, which could be assumed to be placed in the specifier of CP.

(21) [$_{CP}$ [$_{DP}$ Iohannes se godspellere, þe Gode wæs gecweme,] [$_{C'}$ **and**
John the evangelist, who to-God was pleasant and
[$_{IP}$ Crist hine lufode for þære clænnysse]]]
Christ him loved for the cleanness
'St. John, who was pleasant to God, and Christ loved him for his chastity'
(colsigef,ÆLet_5_[Sigefyrth]:26.11)

Furthermore, parallel constructions can be found in subordinate clauses, where a topic is placed in front of a complementizer, as in (22). Constituent fronting to the C-domain that includes an overt C-head may thus be another commonality between CCs and subordinate clauses.

(22) Wite se abbod þeah, [$_{CP}$ [$_{DP}$ eal þæt he do] [$_{C'}$ **þæt**
know.subjunctive the abbot however, all that he does that
[$_{IP}$ he hit do mid Godes ege]]]
he it does with God's awe
'The abbot must know that he should do all that he does in the fear of God'
(cobenrul,BenR:3.16.5.231)

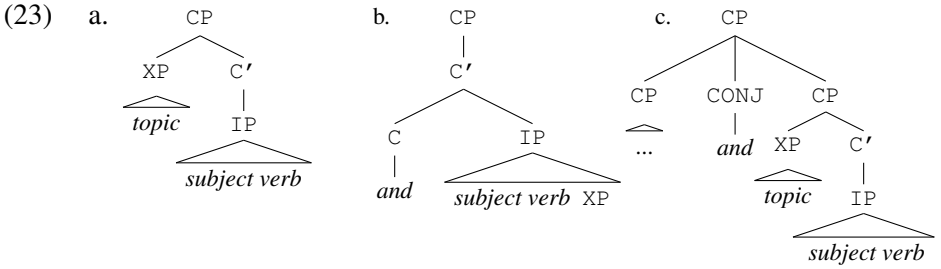
However, structures like (21), for which I found only 5 relevant instances, are considerably rarer than structures like (22), for which there may be about 100 instances in the surviving Old English texts. This difference would be surprising if simultaneous application of topicalization and C-head conjunctions were a genuine grammatical option in Old English. Moreover, alternative parses for the potential counterexamples always remain possible. For instance, (21) could be analyzed with conjunction of the relative clause so that the entire string would be one left-dislocated noun phrase resumed by the pronominal subject of the following sentence.[6] Finally, configurations as in (21) could also result from speech disfluencies, copying errors or Latin interference effects. Thus, it is plausible to maintain the assumption that structures with a topic in the specifier of CP and a conjunction under C are in fact ungrammatical.

[6] I am indebted to Cynthia Allen for pointing out this alternative parse.

4.2 Accounting for the Distributional Difference in Topicalization

The above analysis allows MCs to freely apply the topicalization rule (at least as long as certain requirements imposed by information structure are obeyed), (23a). If a C-head conjunction is present, topicalization will be prohibited, (23b). However, this does not entail that the presence and absence of clause-initial topics distributes in a complementary way between MCs and CCs. Instead, the word order 'conjunction - topic - subject' can still be generated through the usage of logical connectors, (23c). Thus, the proposed analysis successfully models a reduction in, but not a complete lack of, Old English topicalized elements in CCs.

(23)



5 Statistical Evidence II

5.1 Non-Subject Topicalization

I measured the rates of topicalization in MCs and CCs for four non-subject topic categories: non-pronominal object DPs, PPs, ADVPs and ADJPs. The examples in (24) illustrate DPs. The frequencies of four related constructions were determined for each of these topic categories. All of them include subject pronouns as indicators of the position of the IP boundary.

The first structure places the topic XP immediately before the subject pronoun. Additional elements may precede the topic. This configuration indicates topicalization in MCs, (24a). The second structure is parallel to the first except that a conjunction appears in clause-initial position. As before, other elements may precede the topic category. This word order pattern indicates topicalization in CCs, (24b). The third construction type includes the same category that was searched for as a topic but positioned after the finite verb. That is to say, there is a potential topic lower in the structure that could theoretically have fronted but failed to do so (abstracting away from information structure constraints). The clause must be introduced by the subject pronoun. This word order type reflects MCs in which the topicalization rule did not apply, (24c). The last configuration also shows a potential topic category in post-verbal position. Now, however, the clause-initial position must show a conjunction followed immediately by the pronominal subject. Such structures reflect CCs with no topicalization, (24d).

(24) a. [$_{DP}$ Langsume tale] we magon macian be ðysum
long tale we may make about this
'We could write a long tale about this'
(coaelhom,ÆHom_23:80.3745)

b. **&** [$_{DP}$ horses hyda] hi habbað him to hrægle gedon
and horse's hide they have themselves to clothing done
'And they used horse hide for their clothing'
(comarvel,Marv:26.1.130)

c. He arærde ða on ðære ylcan byrig [$_{DP}$ mære cyrcan]
He reared then in the same city great church
'He then built a great church in the same city'
(cocathom2,ÆCHom_II,_38:287.262.6496)

d. **Ac** he worhte [$_{DP}$ fela wundra] ætforan þam deman
but he worked many wonders before the judge
'But he performed many miracles in front of the judge'
(coaelive,ÆLS_[Exalt_of_Cross]:202.5672)

The rate of topicalization can now be calculated as the proportion of sentences with a topic in pre-subject position out of all relevant sentences. Table 3 summarizes the results of this study.

XP category	Clause type	Topicalization *XP - spro - V*	No Topicalization *spro - V - XP*
DP	MC	716 (31.7%)	1,546 (69.7%)
	CC	382 (17.3%)	1,829 (82.7%)
PP	MC	994 (32.9%)	2,024 (67.1%)
	CC	589 (17.5%)	2,768 (82.5%)
ADVP	MC	1,947 (56.7%)	1,484 (43.3%)
	CC	605 (33.1%)	1,222 (66.9%)
ADJP	MC	45 (12.6%)	312 (87.4%)
	CC	12 (3.4%)	337 (96.6%)

Table 3: Topicalization Rates by Category in MCs and CCs

The rate of topicalization is significantly lower in CCs than in MCs for all topic categories. The effect, as measured by the odds ratios, is moderate in all contexts. More precisely, the odds of non-subject topicalization are more than twice as high for MCs than for CCs. The topicalization rates range, for CCs and MCs respectively, from 33.1% and 56.7% for ADVPs, which have the highest propensity to occur before subjects (χ^2=265.6, df=1, p<0.001, odds ratio = 2.65, 95% confidence intervals: 2.35 – 2.98), 17.3% and 31.7% for DPs (χ^2=124.0, df=1, p<0.001, odds ratio = 2.22, 95% confidence intervals: 1.92 – 2.55), 17.3% and 31.7% for

PPs (χ^2=200.9, df=1, p<0.001, odds ratio = 2.30, 95% confidence intervals: 2.05 – 2.59), to only 3.4% and 12.6% for ADJPs (χ^2=18.8, df=1, p<0.001, odds ratio = 4.05, 95% confidence intervals: 2.10 – 7.80). These findings are expected under the assumptions of a model that allows inserting some of the conjunctions in CCs with post-verbal potential topics under C since such a configuration blocks topicalization to the specifier of CP.

5.2 Subject Topicalization

I will now offer one possible measurement of the topicalization rate of non-pronominal subjects in MCs and CCs. Unfortunately, it is not an easy task to ascertain that subjects can be placed in the clause-initial topic position to begin with. The reason is that 'subject - verb' sentences also allow analyses with the subject in the canonical subject position, the specifier of IP.

Nevertheless I assume that parses are available, and in fact common, in which non-pronominal subjects are positioned in the C-domain. Firstly, as shown above, there is ample evidence for the existence of a discourse function in the clausal left periphery through fronted non-subjects and it seems natural to extend the applicability of this position to all grammatical functions, including subjects. Secondly, verbs are virtually always placed under C if they are both negative and subjunctive. Hence, a subject preceding such a verb is probably positioned in the specifier of CP, as in (25).

(25) [$_{CP}$ [$_{DP}$ Ræpsas] [$_{C'}$ **ne** **syn** [$_{IP}$ gesungene mid Alleluian]]]
reply not be.subjunctive sung with hallelujah
'A response (in church service) should not be sung with a hallelujah'
(cobenrul,BenR:15.39.20.518)

Thirdly, non-subject pronouns often occur in the left periphery of the clause in Old English. In this case, they indicate the IP-boundary. A subject preceding such a preposed pronoun is therefore likely to be positioned in the specifier CP. This is illustrated in (26).

(26) [$_{CP}$ [$_{DP}$ min God] [$_{IP}$ **me** asende to _ sona his engel]]
my God me sent to soon his angel
'My God sent his angel to me at once'
(coaelhom,ÆHom_22:326.3470)

Since the frequency of subject topicalization is hard to estimate directly, I used another construction as a proxy. I counted all instances of sentences with left-dislocation and subject resumption. The underlying assumption is that the left-dislocated constituent behaves similarly to a topic in the specifier of CP with the resumptive element occurring within the IP. Left-dislocation and resumption were required to be immediately adjacent so that the resultant structures resemble topicalization more closely. This is illustrated by the parses in (27).

(27) a. [$_{CP}$ [$_{DP}$ se þe ytt þysne hlaf]$_i$ [$_{IP}$ he$_i$ leofað on ecnysse]]
he who eats this loaf he lives in eternity
'He who eats of the bread will live forever'
(cocathom1,ÆCHom_I,_2:192.82.362)

b. **and** [$_{CP}$ [$_{DP}$ se ðe on geleafan wunað]$_i$ [$_{IP}$ se$_i$ leofað on ecnysse]]
and he who in faith lives he lives in eternity
'He who lives in the true faith will live forever'
(coaelive,ÆLS_[Apollinaris]:237.4713)

Next, it is necessary to determine the frequency of clauses in which the topicalization rule fails to apply. I employed for this purpose constructions of the form 'verb - non-pronominal subject' where the verb appears clause-initially. Positive, indicative lexical verbs are normally placed in `I` (although there may be exceptions). Thus, non-pronominal subjects that follow such verbs are positioned in a low subject position, modeled as the specifier of `VP`, and cannot possibly be parsed as topics in the specifier of `CP`. Pronouns or adverbs may intervene between the initial verb and the subject. Subordinate clauses may precede the initial verb. The sentences in (28) exemplify this structure.

(28) a. [$_{IP}$ *Sende* [$_{VP}$ Balthild seo cwen mycel weorod]]
sent Balthild the queen much troop
'Queen Balthild sent a great army'
(cobede,Bede_5:17.456.4.4577)

b. **&** [$_{IP}$ *com* [$_{VP}$ Swegn eorl in mid vii scipon]]
and came Sweyn Earl in with seven ships
'And then Earl Sweyn came in with seven ships'
(Chronicle3_cochronE,ChronE_[Plummer]:1046.21.2188)

The rate of subject topicalization can now be calculated as the proportion of sentences with subject left-dislocations out of all relevant sentences. The result of this investigation is presented below in Table 4.

Clause type	Topicalization *Resumptive Subject - Verb*	No Topicalization *Verb-First - Subject*
MC	583 (79.5%)	150 (20.5%)
CC	276 (64.0%)	155 (36.0%)

Table 4: Subject Topicalization in MCs and CCs

As expected, subjects topicalize significantly less frequently in CCs than in MCs. The association between clause type and topicalization in this sample is comparable to the measurements of non-subject topicalization as shown by a similar value of the odds ratio (χ^2=32.9, df=1, p<0.001, odds ratio = 2.18, 95% con-

fidence intervals: 1.67 – 2.85). Hence, this investigation, too, provides some evidence for a lower rate of topicalization in CCs than in MCs. My model can account for this finding by parsing some of the clause-initial conjunctions in verb-first structures as C-head conjunctions thus blocking the C-domain as a potential locus for the subject.

6 Conclusion

In this paper, I argued that a syntactic model in which conjunctions may vie for a structural slot with finite verbs successfully captures the distributional differences regarding verb placement and constituent fronting between Old English main clauses and main clauses introduced by a conjunction. Various extensions of this work are conceivable. Firstly, the findings could be strengthened by investigating the word order distributions in more detail, for example for different conjunctions, by individual texts or in specific information-structural alignments. Secondly, it may be possible to discover semantic differences between C-head conjunctions and logical connectors. Finally, and most importantly, the diachronic development of word order patterns in MCs and CCs should be investigated since the postulation of C-head conjunctions has interesting theoretical implications for the predicted trajectory of change for the loss of head-final IPs and high verb placement.

A Text Files Used for the Quantitative Studies

The electronic text files listed below were used for the quantitative studies. The YCOE manual offers details on each text file, such as word count and underlying edition. The texts do not all come from exactly the same period, but were composed roughly between 875 - 1100.

coalex.o23.psd, comarvel.o23.psd, cochad.o24.psd, cochristoph.psd, cosolsat2.psd, comart1.psd, comart2.psd, comart3.o23.psd, cochronA.o23.psd, cochronC.psd, cochronD.psd, cochronE.o34.psd (duplicate content in the Anglo-Saxon Chronicle files [A]-[E] has been removed), coprefcura.o2.psd, cocura.o2.psd, cocuraC.psd, cobede.o2.psd, coboeth.o2.psd, coorosiu.o2.psd, coprefsolilo.psd, cosolilo.psd, codocu1.o1.psd, codocu2.o12.psd, codocu2.o2.psd, codocu3.o23.psd, codocu3.o3.psd, codocu4.o24.psd, colaece.o2.psd, coherbar.psd, coquadru.o23.psd, conicodA.psd, covinsal.psd, coblick.o23.psd, coverhom.psd, coverhomE.psd, coverhomL.psd, cobenrul.o3.psd, cochdrul.psd, comary.psd, coeuphr.psd, cosevensl.psd, cowsgosp.o3.psd, codicts.o34.psd, coprefcath1.o3.psd, cocathom1.o3.psd, coprefcath2.o3.psd, cocathom2.o3.psd, coaelhom.o3.psd, copreflives.o3.psd, coaelive.o3.psd, coprefgen.o3.psd, coepigen.o3.psd, cootest.o3.psd, cotempo.o3.psd, colwsigeXa.o34.psd, colwgeat.psd, colsigewZ.o34.psd, colsigef.o3.psd, colwstan1.o3.psd, colwstan2.o3.psd, coalcuin.psd, coinspolX.psd, cocanedgX.psd, cowulf.o34.psd, colaw6atr.o3.psd (the Laws of Aethelred VI are the only laws included because they were written by Wulfstan), cobyrhtf.o3.psd, colacnu.o23.psd, coapollo.o3.psd, conicodD.psd, corood.psd, cojames.psd, coeust.psd, comargaT.psd, comargaC.o34.psd, coleofri.o4.psd, coneot.psd

References

Bech, Kristin. 2001. *Word Order Patterns in Old and Middle English: A Syntactic and Pragmatic Study*. PhD Dissertation: University of Bergen.

Besten, Hans den. 1983. On the Interaction of Root Transformations and Lexical Deletive Rules. In Werner Abraham (ed.) *On the Formal Syntax of the Westgermania*, pages 47–131. Amsterdam: John Benjamins.

Kemenade, Ans van. 1987. *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

Kemenade, Ans van. 1997. V2 and Embedded Topicalisation in Old and Middle English. In Ans van Kemenade and Nigel Vincent (eds.) *Parameters of Morpho-Syntactic Change*, pages 326–352. Cambridge: Cambridge University Press.

Light, Caitlin. 2012. *The Syntax and Pragmatics of Fronting in Germanic*. PhD Dissertation: University of Pennsylvania.

Mitchell, Bruce. 1985. *Old English Syntax*. Oxford: Clarendon.

Pintzuk, Susan. 1999. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. New York: Garland.

Speyer, Augustin. 2010. *Topicalization and Stress Clash Avoidance in the History of English*. Berlin: de Gruyter.

Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. *The York-Toronto-Helsinki Corpus of Old English Prose (YCOE)*. http://www-users.york.ac.uk/ lang22/YCOE/YcoeHome.htm (Accessed 10 April 2014). Oxford Text Archive.