The Proleptic Accusative as an exceptional Exceptional Case Marking construction
Lucas Seuren, Marjo van Koppen & Mark de Vries
University of Groningen and Utrecht University
– version June 2014 –

Abstract. This paper presents a detailed analysis of the so-called proleptic accusative. This construction type is rare, but attested in a number of different languages. Here, we discuss data from Middle Dutch in particular, which enables us to shed new light on the matter. We show that the proleptic accusative is crucially different from complex cases of raising. Instead, we explore two novel and competing hypotheses about its structural properties. We first argue that an analysis in terms of clausal coordination and ellipsis straightforwardly solves a number of syntactic complications, but this leads to problems with semantic interpretation in various cases. We then propose that the seemingly additional accusative argument in the matrix can be base-generated as an embedded hanging topic of the complement clause involved. This requires exceptional accusative case marking across a clause boundary by the matrix verb. The necessary combination of various prerequisites explains why the proleptic accusative is only sporadically found across languages. Importantly, however, it is now clear how this complex construction type can be decomposed into more basic syntactic ingredients, voiding the need to enrich the model of grammar with additional stipulations.

Keywords: prolepsis, exceptional case marking (ECM), hanging topic left dislocation (HTLD), embedded topicalization, clausal ellipsis, copy raising, resumptive pronouns, aboutness

1. Introduction

In this article, we provide a detailed analysis of the so-called proleptic accusative. As we will show, the apparently contradictory properties of this construction type provide an interesting challenge to the theory of grammar. Although the proleptic accusative appears to be cross-linguistically rare, it is attested in different and even unrelated languages, including classical Latin and Greek, and Nahuatl. Here, we present data from Middle Dutch (making use of a small corpus established by Stoett 1923) as well as some novel data from present-day German, which shed new light on the matter. Some of the advantages of looking at continental Germanic, is that these languages do not have pro-drop, contrary to the other ones mentioned, and that they have a slightly more fixed canonical word order, which makes it easier to determine the syntactic status of the constituents involved.

A prototypical example of the construction at hand is provided in (1), where we use English words for ease of representation, postponing actual data and structural variations and complications of (1) to the sections below.

(1) We saw this man [that he came in peace].
‘About this man, we saw that he came in peace.’

The underlined constituent this man is the relevant accusative argument, taken up by the bold-faced resumptive pronoun he in the embedded clause between brackets. What is remarkable here is that this man appears to be part of the matrix clause, whereas it is semantically related to the contents of the complement clause following it. Within this
embedded clause, the resumptive pronoun is used as a pro-form of the accusative noun phrase. According to this assessment – reflected by the term prolepsis ‘anticipate’ from ancient Greek, suggesting that something is realized unexpectedly early in the utterance – (1) might as well be formulated more straightforwardly as in (2):

(2) We saw [that this man came in peace].

Nevertheless, notice right away that the idiomatic translation in (1) deviates from (2) in that it adds an aboutness meaning aspect to the sentence (we return to this issue).

Setting aside details of the interpretation for a moment, we immediately perceive a number of remarkable syntactic problems in (1), which need to be resolved. These are absent, by the way, in periphrastic aboutness/prolepsis constructions, as we show later. The main issue is that there appear to be two constituents eligible as the direct object of the matrix verb saw, namely the accusative noun phrase this man as well as the embedded clause that he came in peace. This raises questions about the status of either constituent: which one is the real internal argument of the verb, and then what is the other one? If the clause is a complement clause, where does the noun phrase find its origin and theta role, and where exactly does it surface? And how does it acquire accusative case?

At first blush, one might think that the noun phrase is generated in the subordinate clause and raised to the matrix, corresponding to the situation in regular raising constructions as in (3), where the theta role is provided in the base position and Case in the surface position.

(3) This man, seems [t, to come in peace].

However, this line of analysis is problematic for the proleptic accusative in (1), as it would require raising-to-object across a finite CP boundary; moreover, it leaves the question where the nominative resumptive pronoun comes from.

We will advance and investigate two alternative views on the structure of the proleptic accusative. We first argue that an analysis in terms of clausal coordination and ellipsis straightforwardly solves a number of the syntactic complications mentioned. The basic idea is illustrated in (4), where both the noun phrase and the complement clause are treated as the direct object of the verb saw. This is possible by simple doubling the matrix clause in a juxtaposed configuration, and apply some form of clausal ellipsis or gapping in the second conjunct.

(4) We saw this man; we saw that he came in peace.

Since this leads to problems with semantic interpretation in various cases, we also propose an entirely different alternative. This analysis is syntactically more complicated, but appears to be more faithful to the perceived meaning of the construction. On this view, the seemingly additional accusative argument in the matrix is base-generated as an embedded hanging topic of the complement clause, as sketched in (5):

(5) We saw [ [this man], [that he came in peace]].

This requires exceptional accusative case marking of this man by the matrix verb across the finite clause boundary.
Thus, we relate the proleptic accusative to aboutness left dislocation, embedded topicalization phenomena as well as exceptional case marking, illustrated in (6a/b/c), respectively:

(6) a. [This man, we gave him a book].
    b. We think [that this man, we need him for the job].
    c. We heard [him give a talk].

It seems to us that the rather uncommon combination of structural mechanisms is the reason that the proleptic accusative is only rarely found across languages. What is crucial, however, is that there is no need to stipulate construction-specific properties. In line with the general tenets of generative grammar, and in particular the minimalist program, we can decompose what appear to be complexities of the construction type at the level of human meta-analysis to the interaction of more basic functionalities of the syntactic system.

The article is set up as follows. In section 2, we elaborate on the Middle Dutch data, and show in more detail why analyses in terms of raising are inadequate. In section 3, we introduce the possibility of a biclausal structure accompanied by ellipsis. In section 4, we develop the alternative in terms of embedded dislocation, and offer some cross-linguistic considerations. Section 5 is the conclusion.

2. Syntactic properties of prolepsis

2.1. Overview of the data

The proleptic accusative construction is characterized by the presence of two constituents that may function as the direct object of the matrix verb: a finite complement clause and a noun phrase that is accusative-marked (whether morphologically visible or abstractly). Furthermore, the embedded clause contains an argument, normally a resumptive pronoun, that is interpreted as coreferent with the relevant preceding noun phrase. In Middle Dutch (~1200-1500 A.D.), we clearly see this pattern:

(7) Maer die serjanten zijn kenden den coninc van Israël, dat hi niet was harde fel.
    lit. ‘But his sergeants knew the king of Israel, that he wasn’t very fierce.’
    (Rijmbijbel, v.12643, translation ours)

In (7), the proleptic accusative noun phrase is den coninc van Israël ‘the king of Israel’; the resumptive pronoun in the embedded clause is the nominative hi ‘he’. Generally, the verb kennen ‘know’ can take nominal or clausal complements.\(^2\)

The proleptic accusative construction is related to, but not the same as periphrastic equivalents containing a sentence-initial or parenthetically intercalated prepositional

---

1 For clarities’ sake, we use simple word translations in the glosses, only marking the most relevant grammatical features explicitly.

2 Note that the Modern Dutch verb kennen ‘be acquainted with’ is semantically more specialized than its Middle Dutch counterpart; it no longer includes certain meaning aspects of weten ‘have knowledge of’.
aboutness phrase *about/concerning/as for... DP*. An illustration from present-day Dutch is provided in (8):

(8) [Wat betreft de koning], ik denk dat hij binnenkort zal aftreden.
what concerns the king I think that he soon will abdicate
‘As for the king, I think he will abdicate soon.’

Such independent aboutness phrases do not display all the intriguing properties of the proleptic accusative discussed here.³

Let us return to (7), then, and ask the following crucial question: What is the empirical range of variation? At least four aspects of the proleptic accusative construction seem relevant: (i) the linear position of the proleptic ‘object’ in the matrix clause, (ii) the syntactic role and form of the resumptive element, (iii) the nature of the embedded clause, and (iv) the nature of the verbs involved. We discuss and illustrate these in turn.

As we have seen, the ACC-DP (or ‘proleptic object’) may occupy the position after the matrix verb, immediately preceding the embedded clause. In an embedded environment it would require Heavy NP Shift – in a descriptive sense – to obtain the same linear configuration, considering that Dutch is superficially OV, and has verb second (V2) in main clauses but not subordinate clauses. Example (9) illustrates this pattern:

(9) ... als hy verstond den helighen man dat hi sceeden wilde
when he understood the ACC holy ACC man that he depart wanted
van dan.
from there

lit. ‘...when he understood the holy man, that he wanted to leave.’

(Leven van Sinte Amand II, v.4676-4677, translation ours)

In addition, there are examples in which the noun phrase precedes the final verbal position (the ‘right sentence bracket’); see also below. Here, we illustrate that the ACC-DP can be a clitic in the higher middle field:

(10) ... waer si-ne vernam dat hi in enighe lande quam.
where she-him ACC heard that he in some countries came

‘...where she heard him, that he came in some countries.’

(Der minnen loep, v.873-874, translation ours)

Even more interesting is the fact that the ACC-DP can be topicalized in the matrix, which is in fact a more common configuration; see (11):

³ There is also an integrated – adverbial – variant with *of* PPs, which has a more limited distribution.

(i) [Van welk boek] denk je dat alle studenten het interessant vinden?
of which book think you that all students it interesting find

lit. ‘Of which book do you think all students find it interesting?’

Salzmann (2006) discusses this type of ‘resumptive prolepsis’ in the context of relative clauses (e.g., ... *the book of which the students say they find it interesting*).
(11) **Desen Tyberius hor-ic liën dat hi tien tiden ontboot menegen**

this\textsubscript{ACC} Tiberius hear-I tell that he that time summoned many

coninc ende heren groot.

king and lord big

\textit{lit.} ‘I’ve heard Tiberius, that in that time he summoned many great kings and lords.’

(Spieghel Historiael I: VII-III, v.36-38, translation ours)

In each of the examples above, the resumptive pronoun is the nominative-marked subject of the embedded clause. This is the most common situation, but not the only one. In (12), the resumptive element is the object of the lowest clause:

(12) Der Walewein claechde \textit{sijn swert} dat hij \textit{-t} daer niet \textit{en} hevet.

the Walewein complained his sword that he-it there not \textit{NEG} has

\textit{lit.} ‘Walewein complained his sword, that he didn’t have it there.’

(Roman van Walewein, v.8125-8127, translation ours)

A further example worth highlighting at this point is (13), in which the ACC-DP is a coordinated phrase, and there are two resumptive pronouns, the subject and object of the embedded clause.

(13) Wi vinden \textit{oec in den nieuwen testamente Christum} ende \textit{sine} apostele, we find also in the new testament \textit{Christ} and his apostles

\textit{dat si} \textit{hem} keerden ende overgaven \textit{in die doot}.

that they\textit{NOM} him\textit{ACC} turned and \textit{over.gave} \textit{in the death}

\textit{lit.} ‘In the new testament, we also find Christ and the apostles, that they gave him over in death.’

(Vanden Gheesteliken Tabernakel, p.19, translation ours)

We will return to this phenomenon below.

Another aspect of the proleptic construction that varies is the type of embedded clause involved. According to Verdam (1908), it is a complement clause introduced by \textit{dat} ‘that’, as in the examples shown hitherto. However, we also found embedded clauses introduced by a \textit{wh}-phrase, i.e., embedded questions; see (14), for instance:

(14) Den \textit{inghel} horden wi nochtan \textit{hoe hi} sprac den wiven an.

\textit{the\textsubscript{ACC} angel} heard we still \textit{how he} spoke the women to

\textit{lit.} ‘The angel, we still heard how he talked to the women.’

(Spieghel Historiael I: VII - XXXVI, v.17-18, translation ours)

Like (11) and other examples, (14) has the prototypical features of prolepsis, with both an ACC-marked noun phrase and a complement clause that seem to compete for the function of direct object of the matrix verb. The noun phrase also co-refers with a resumptive pronoun in the embedded clause. Another relevant illustration is (15), where the ACC-DP is not topicalized, but figures in the middle field. Note that the matrix clause must be an embedded clause itself, here, considering that it is verb-final.
Finally, let us briefly address the lexical properties of the matrix verb. A question that needs to be answered is whether prolepsis is possible with just any transitive verb. This does not seem to be the case. Although one cannot perform judgment tasks on extinct languages, our impression is that the relevant verbs are typically verbs of perception (‘see’, ‘hear’) or cognition (‘think’, ‘know’). It is no coincidence, we believe, that these are the kind of verbs that cross-linguistically license Exceptional Case Marking (or *accusativus cum infinitivo*), illustrated for Middle Dutch in (16).

(16) a. Die coninghinne hevet vernomen haren heren den coninc comen.
    the queen has heard her husband the king come
    ‘The queen heard her husband, the king, come.’

   b. Si vinden di te wesen een haven des vreden.
    they find you to be a haven the peace
    ‘They consider you to be a haven of peace.’

(Stoett 1923:136-137, translation ours)

Here, the thematic external arguments of the embedded clauses are assigned ACC Case by the matrix verb, which takes an infinitival clause as its complement.

We return to the similarities between ECM and prolepsis in section 4. It is to be noted, however, that there are also clear differences. Most importantly, the embedded clause in the prolepsis construction is finite rather than infinite (also, the CP domain is lexically filled), and there is a Case-marked resumptive pronoun, which is lacking in the ECM construction. Thus, prolepsis cannot simply be equated with regular ECM.

Before we discuss further analytical options, let us make a succinct cross-linguistic comparison. As indicated in the introduction, the proleptic accusative has been reported for a small number of languages, including Latin, Classical Greek, and Nahuatl. In section 4, we will also introduce some new data from present-day German. Now, consider the following sentences from Classical Greek and Latin, respectively:

(17) Katamathete ta krina tou agrou poos auxanousin.
    consider the lilies:ACC the field how grow:3PL
    lit. ‘Consider the lilies in the field, how they grow.’

(Fraser 2001, translation ours)

(18) Lebsonicum hic adulescentem quaero in his regionibus ubi habitet.
    L.:ACC here adolescent:ACC ask:1SG in these:ABL regions:ABL where live:3SG
    lit. ‘Lesbonicus, a young man from these regions, I am asking where he lives.’
    D. Gary Miller (p.c., translation ours)

As in Middle Dutch, we see an accusative noun phrase in addition to a finite complement clause. The embedded clause happens to be an embedded question here, comparable to the
situation in (14) and (15).\(^4\) We do not actually see a resumptive pronoun in the embedded clause, but we can safely assume that it is abstractly there, since (i) Greek and Latin are pro-drop languages (which is related to rich agreement on the verb), and (ii), the clause is finite.

Similarly, a prolepsis pattern is described for Nahuatl, which is a full pro-drop language, with both subject and object marking on verbs. We return to this example in the next section.

(19) \text{niki:nkayik i:n ta:kah (ke) wa:lankeh.} \hfill \text{(Higgins 1981)}

\text{lit. ‘I heard the men, that they come.’} \hfill \text{(literal translation ours)}

Since pro-drop was highly marked in Middle Dutch, one expects an overt resumptive pronoun to show up, which is indeed what we found so far. But it is not completely impossible (cf. Van Helten 1883), and there is indeed one example that can be explained if we assume pro-drop, similarly to the situation in the languages just reviewed:

(20) \text{... ende doe hi there ons Heren sach, hoe dat in die tenten lach.} \hfill \text{(Rijmbijbel, v.6087-6088, translation ours)}

\text{lit. ‘... and when he saw our lord there, how he lay in the tent.’}

In (20), an overt subject of the finite embedded clause is lacking; the silent pronoun is interpreted as coreferent with the object in the matrix (\textit{ons Heren} ‘our lord’).

To wrap up the discussion so far, the properties of the prolepsis construction are summarized schematically in (21):

(21) \text{Properties of the proleptic accusative construction:}

\begin{itemize}
  \item \text{a. the linear position of the accusative/object DP:}
    \begin{itemize}
      \item (i) immediately preceding the embedded clause, and after the verb, or
      \item (ii) topicalized, or
      \item (iii) in the middle field;
    \end{itemize}
  \item \text{b. the nature of the complement clause:}
    \begin{itemize}
      \item (i) finite \textit{that}-clause, or
      \item (ii) finite embedded \textit{wh}-question;
    \end{itemize}
  \item \text{c. the syntactic function of the resumptive element:}
    \begin{itemize}
      \item (i) subject (usually), or
      \item (ii) object ;
    \end{itemize}
  \item \text{d. the lexical nature of the resumptive element:}
    \begin{itemize}
      \item (i) overt personal pronoun, or
      \item (ii) covert \textit{pro}, if pro-drop is allowed;
    \end{itemize}
  \item \text{e. the lexical nature of the matrix verb:}
    \begin{itemize}
      \item usually verbs of perception or cognition (coinciding with ECM verb types)
    \end{itemize}
\end{itemize}

\(^4\) According to Fraser (2001), prolepsis in Classical Greek often involves a \textit{wh}-clause, whereas we saw that in Middle Dutch \textit{that}-clauses appear to be more frequent. Since such differences appear to be only tendencies in the frequency of use, and do not correspond to absolute differences in grammaticality, we do not consider them relevant for our analysis.
2.2. Towards an analysis

With the empirical foundation settled, let us turn to potential analyses of prolepsis. Although the literature is scarce, a number of proposals – mostly sketchy or implied – have been made to analyze prolepsis. They roughly fall into one of two categories: raising or base-generation. That is, either the proleptic object is generated in the embedded clause and raises to the matrix clause, leaving behind a pronoun (Stoett 1889, Van Gestel et al. 1992), or it is generated as an additional argument in the matrix clause (Higgins 1981, Maraldi 1986, Ura 1984).

The idea of raising (more specifically, ‘copy hyper-raising to object’) is sketched in (22). What it solves is the issue of theta roles: the proleptic object is thematically related to an argument position in the embedded clause, and this is now where it is generated.

\[
\begin{align*}
\text{a. } & [\text{matrix } \ldots V [\text{emb. clause } \ldots \text{DP}_{\text{prol.}} \ldots]] \Rightarrow \\
\text{b. } & [\text{matrix } \ldots V \text{DP}_{\text{prol.}} [\text{emb. clause } \ldots \text{res.pron } \ldots]]
\end{align*}
\]

But difficult questions need to be addressed: (i) where does the resumptive pronoun come from? It must be some kind of spelled-out trace; (ii) how can case assignment to the proleptic constituent in the higher clause overrule or prevent case assignment in the embedded clause?; (iii) what triggers movement, and what are the features and positions involved? It might be that the pronoun is the stranded head of a ‘Big DP’-structure (compare independent proposals by Boeckx 2003, among others). But even if technical solutions can be found, serious problems remain. For instance, we have seen that the embedded clause can be a wh-clause; recall (14) and (15). This would in principle make extraction from the embedded position ungrammatical for the same reason that (23) is out.

\[
\begin{align*}
\text{(23) } & * \text{ Which man, did you hear [how } ti \text{ talked to the woman].}
\end{align*}
\]

Furthermore, there does not seem to be any plausible way to derive the example involving constituent coordination in (13) in this way. We would either need to move a non-constituent, or lower a moved phrase into a coordination phrase, neither of which is possible.

Thus, we discard an analysis in terms of raising. Simple base-generation is not really an option either, because it is merely a restatement of the facts. That is, we want to refrain from postulating prolepsis as a construction in which the matrix verb exceptionally takes two internal arguments instead of one, while this is otherwise impossible. Notice, moreover, that if the proleptic argument is not resumed in the embedded clause, the construction is intuitively unacceptable (setting aside potential elaborations of the Loose Aboutness type, as in Speaking about cancer, John doesn’t smoke). Compare for instance the nonsensical attempts in (24a/b) to (7) and (14), respectively:

\[
\begin{align*}
\text{a. } & * \text{ But his sergeants knew the king of Israel, that Mary wasn’t very fierce.} \\
\text{b. } & * \text{ The angel, we heard how John talked to the women.}
\end{align*}
\]

There is, however, no straightforward way to exclude the derivation of such sentences from a simple base-generation analysis.

Having said that, we immediately want to point out that it is possible to improve on the basic ideas just sketched (see also Massam 1985 for discussion). In what follows, we investigate potential solutions involving base-generation in different ways, as well as the possibility of movement being involved, though not of the raising type.
3. Prolepsis as ellipsis?

As we saw in the previous section, the core characteristic of the prolepsis construction is the presence of an additional argument for which a structural position seems to be lacking. This puzzling property is clearly a challenge for any explanation. In an attempt to resolve the issue, let us explore a direction not previously entertained in the literature, as far as we are aware.

The accusative noun phrase and the embedded clause involved can be seen as competitors for the object role in the matrix. From this, it may be concluded that in fact two available argument positions are necessary. Needless to say, we want to avoid stipulating that the relevant verb licenses an additional argument in this particular construction only. There is, however, an alternative and at first sight more plausible solution, which reverses the perspective: if there are two internal arguments, there must be two verbs. This can be effectuated by means of ellipsis: the two verbs are lexically the same, and for this reason only one needs to be spelled out. Thus, consider the bisentential (or biclausal) analysis sketched in (20), adorned with the hypothetical example repeated from (1):

(20) a. \[ s_1 \text{ SU V DO}_\text{DP}; s_2 \text{ (SU V) DO}_\text{CP} \].

b. We saw this man; (we saw) that he came in peace.

There are two juxtaposed matrix clauses, abbreviated S1 and S2 here. In the first, the accusative noun phrase functions as the object, and in the second, the embedded complement clause does. Since the subject and the matrix verb remain constant, they can be deleted in sentence 2. The embedded clause survives as an ellipsis remnant.

From a syntactic viewpoint, it is reasonable to analyze both the DP and the CP as objects of the matrix verb. As was commemorated in the previous section, all verbs used in prolepsis constructions are in fact verbs that license both nominal and clausal complements independently. Below, we show that this solution straightforwardly accounts for the other syntactic properties of the prolepsis construction. But let us first add a few remarks on the overall configuration and the deletion at hand.

The proposal in (20) relates prolepsis to other instances of clausal ellipsis. Relevant examples include sluicing, fragment answers, and afterthoughts (see Merchant 2001, Ott & De Vries to appear a/b, among many others):

(21) a. Peter bought something, but I don’t know what Peter bought.

b. A: What did Peter buy?
   B: A book about linguistics Peter bought.

c. Peter bought something interesting yesterday: a book about linguistics Peter bought yesterday.

It is usually assumed that the remnant is fronted (by means of A’-movement) in the elliptical clause.\(^5\) The given information following the focused remnant is phonologically deleted, but there is clear evidence that it is syntactically present (cf. Van Craenenbroeck & Merchant 2013 for discussion). Licensing of ellipsis as such is essentially semantic, but may go hand in hand with syntactic feature checking (see, for instance, Aelbrecht 2010 on ‘e-givenness’, and Thoms 2010 for a different take on the issue).

---

\(^5\) Extended gapping around *in situ* remnants is a feasible alternative, but this has non-constituent deletion as a potential disadvantage. Such issues are tangential to the purpose of this paper.
Here, we do not have anything to add about the general theory of ellipsis; it suffices to
note that the usual procedure with all the necessary constraints can simply be applied to a
bisentential analysis of prolepsis as well; nothing new needs to be postulated. In (20), the two
main clauses are completely parallel, apart from the internal structure of the focused objects,
of course. After fronting of the complement clause in the second matrix clause, ellipsis can
take place:

(22) ... \[S_2 \text{DO}_C \text{P} \text{SU} \rightarrow \text{CA}\].

How the two juxtaposed sentences are syntactically related to each other is a separate
question. One can easily imagine a structure in terms of (silent) coordination. As is clear from
(21), overt coordination is not a formal prerequisite for clausal ellipsis. The only condition
seems to be that the two clauses are adjacent to each other in discourse, and in some way
pragmatically related. This is evidently the case in (20).

Let us now turn to some concrete examples, and see in more detail how the approach
fares. In (23), the verb ontrieden ‘feared’ takes the noun phrase den swerten here ‘the black
lord’ as its object in the first sentence. The second sentence can be understood as an
elaboration or specification of the first. In S2, the verb takes a complement clause, dat hi-ne
soude te doet slaen ‘that he would kill him’. This focused clause is fronted, and the verb and
subject, which are e-given in S2, get deleted. For concreteness’ sake, we applied verb second
(V2) in S2, but strictly speaking this is irrelevant in an elliptical context.

(23) a. Si ontrieden den swerten here dat hi-ne soude te doet slaen.
they feared the ACC black ACC lord that he-him would to death beat
lit. ‘They feared the black lord, that he would kill him.’
(Ferguut, v.1724-1725, translation ours)

b. \[S_1 \text{ Si ontrieden den swerten here}; [S_2 [dat hi-ne soude te doet slaen], ontrieden
\[\text{ Si-t}]]

We expect both sentences to have a full-fledged argument structure, and to be internally
complete in a syntactic sense. The proleptic accusative object is den swerten here ‘the black
lord’, which is visibly case-marked in S1. The resumptive pronoun is hi ‘he’, which is now
the regular nominative subject of S2. There is no syntactic link between these two elements,
and coreference is simply an instance of cross-sentential anaphora, similar to the situation in
(24):

(24) They feared the black lord. He, might kill someone.

In sum, the distribution of arguments and Case in (23) seems straightforward (but see below).

For examples like (25), we can say basically the same, with the proviso that the relevant
juxtaposed clauses are not main clauses, but subordinated themselves (in this case, introduced
by the complementizer als ‘dan’).

(25) a. ... als hy verstand \[den helighen man dat hi sceeden wilde
when he understood the ACC holy ACC man that he depart wanted
van dan.
from there
lit. ‘...when he understood the holy man, that he wanted to leave.’
(Leven van Sinte Amand II, v.4676-4677, translation ours)
b. \( [S_1 \text{ dat hy verstond [den helighen man]}] [S_2 \text{ [dat hi sceeden wilde van dan], verstand hy}] \)

It is important to see that clausal ellipsis in embedded environments is more generally allowed; see the examples in (26), for instance. Here, we take the view that in the relevant cases, ellipsis can directly target embedded clauses, rather than repair islands, as explicated in De Vries (2013) for Right Dislocation (see also Ott & De Vries to appear b).

(26) a. I heard that Peter bought something but Mary doesn’t know what \textit{he} bought.
b. Because Peter bought something and I didn’t know what \textit{he} bought, I asked what it was.
b. I talked to a man who bought something interesting today: a book about linguistics \textit{he} bought.

Without going into detail, we simply observe that the proleptic accusative construction fits this general pattern.

Next, recall that the embedded object clause can be a \textit{wh}-clause. Furthermore, the accusative noun phrase can be topicalized within the matrix. We show both properties at once in (27).

(27) a. \textit{Den ingel horden wi nochtan hoe hi sprac den wiven an.} the\textsubscript{acc} angel heard we still how he\textsubscript{nom} spoke the women to
\textit{lit.} ‘The angel, we still heard how he talked to the women.’

\textit{(Spieghel Historiael I: VII - XXXVI, v.17-18, translation ours)}
b. \( [S_1 \text{ \textit{Den ingel} horden wi nochtan \( t_i \)}]; [S_2 \text{ [hoe hi sprac den wiven an], horden wi nochtan \( t_i \)}]. \)

The \textit{wh}-character of the embedded clause is unproblematic: complement clauses can be embedded questions, and in this analysis nothing needs to be extracted out of it. In S1, the accusative object is topicalized. This too, is generally very common in Dutch, so it comes as no surprise that it is possible in a prolepsis construction.

Since the relationship between the proleptic accusative and the resumptive pronoun is only one of cross-sentential coreference on the current view, there does not seem to be a compelling reason why the resumptive pronoun should be the subject of the embedded clause. Indeed, examples are attested in which it is an object; see (28):

(28) a. \textit{Der Walewein claechde \textit{sjin swert} dat hij-t daer niet en hevet.} the Walewein complained his sword that he-it there not NEG has
\textit{lit.} ‘Walewein complained [about] his sword, that he didn’t have it there.’

\textit{(Roman van Walewein, v.8125-8127, translation ours)}
b. \( [S_1 \text{ Der Walewein claechde \textit{sjin swert}]; [S_2 \text{ [dat hij-t daer niet en hevet], claechde der Walewein} \( t_i \)]. \)

An interesting example, impossible to explain under a raising account, is repeated in (29a). Here, the proleptic accusative related to the nominative resumptive pronoun is embedded in a coordinated phrase. As shown in (29b), this is no longer a problem in a bisentential analysis.
(29) a. Wi vinden oec in den nieuwen testamente Christum ende sine apostele, we find also in the new testament Christ and his apostles
dat si hem keerden ende overgaven in die doot.
that they him turned and over.gave in the death
lit. ‘In the new testament, we also find [about] Christ and the apostles, that they
gave him over in death.’
(Vanden Gheesteliken Tabernakel, p.19, translation ours)
b. [S₁ Wi vinden oec in den nieuwen testamente [Christum ende sine apostele]]; [S₂ dat si hem keerden ende overgaven in die doot], vinden wi oec in den nieuwen testamente t.

Notice, however, that contrary to the traditional view, we now know that resumption can also involve a pronominal object, as was just illustrated in (28). If so, (29) probably involves resumption of both the subject and the object inside the embedded clause. The accusative antecedent, then, is not a subconstituent of the coordinated phrase, but the entire phrase, as shown in (30):

(30) Wi vinden... Christum and sine apostele, dat si hem keerden...

We can illustrate the same phenomenon in English. In (31a), for instance, the coordinated object of the first sentence can be resumed by two pronouns in the next sentence, which is arguably about both Peter and Anne.

(31) a. I saw Peter and Anne. She kissed him passionately.
b. I looked at the two lovers. One/Anne started kissing the other/Peter.

This general process of split pronominalization or split referencing is the reverse of the more familiar split antecedent taking, as in (32).

(32) Peter met Anne in the supermarket. They/P.&A. decided to walk home together.

Since such possibilities are independent properties of the language system, they should not (and on our analysis do not) require a separate explanation within the context of the prolepsis construction.

A final example we like to highlight concerns the object agreement pattern in Nahuatl. In (33), there is a 3PL object marking on the matrix verb niki:nkayik ‘heard’, triggered by the plural object i:n ta:kah ‘the men’.

(33) niki:nkayik i:n ta:kah (ke) wa:lankeh.
1SG.3PL.hear.PRET the men (COMP) 3.come.PRET.PL
‘I heard the men come.’
lit. ‘I heard the men, that they come.’
(Higgins, 1981)

In the absence of the proleptic object, however, the verb form would be nikayik, with default SG agreement for the complement clause. This gets a natural explanation on the bisentential analysis, in which the overt matrix verb agrees with the plural nominal object in S₁, and the elliptical verb with the default singular form is in S₂. This is shown in (34):

(34) [S₁ niki:nkayik i:n ta:kah]; [S₂ [(ke) wa:lankeh] nikayik t].
More generally, ellipsis is insensitive to number agreement; see (35) in English, for instance. Here, the verb form *likes* takes the verb *like* as a licit antecedent.

(35) a. The boys like ice-cream and Anne (does), too.
   b. [S\textsubscript{1} The boys like ice-cream] and [S\textsubscript{2} Anne *likes* ice-cream, too].

This follows from what we mentioned earlier, namely that ellipsis licensing is essentially semantic.

All in all, it seems that an analysis in terms of ellipsis explains the full set of syntactic properties of the prolepsis construction. However, this should not conceal the fact that there are problems of interpretation, which we discuss now.

Consider (23) again – *lit.* ‘They feared the black lord, that he would kill him.’ Here, the subjects feared that the black lord would kill a person. It is not implausible that they did not only fear such an event, but also the black lord himself. It might be, therefore, that the noun phrase is thematically related to the matrix verb. However, it is also feasible that fear of the black lord himself is merely a potential inference, and not to be directly encoded in the argument structure. If that is the case, the analysis presented is inadequate.

A paraphrase of the intended meaning could be that the subjects feared *about* the black lord that he would kill someone. There are other examples in which the aboutness meaning aspect is more compelling. Take (36), for example (and see also (28) above). Here, the subject does not command the relevant statues themselves, but he orders about those statues that they be carried together.

(36) So heeft-i geboden die beelden van-den afgoden dat men-se te samen drouge sciere.
     *lit.* ‘Thus he commanded [about] the statues of the false gods, that people carry them together.’
     (Spieghel Historiael III: XXXVIII, v. 21-22, translation ours)

Therefore, it is semantically odd to generate the statues as the internal argument of the matrix verb, even though it is syntactically fine.

The problem at hand is parallel to the situation in exceptional case marking (ECM) constructions. Consider (37a/b):

(37) a. I heard Peter fall down the stairs.
   b. I found him gone.

What the subject hears in (37a) is not Peter himself, but the sound that the falling event makes; in (37b), the subject does in fact *not* find him, but finds out that he is gone. The standard analysis is therefore that the internal argument of the matrix verb corresponds to a (small) clausal complement that originally includes the accusative-marked noun phrase. This phrase thus receives its theta-role but not its case-marking in the embedded clause.

In the next section we take the parallel between prolepsis constructions and ECM further. The challenge is how we can maintain the advantage of the ellipsis analysis that the proleptic accusative is generated independently of the resumptive pronoun, and still prevent it from being thematically related to the matrix verb.
4. Prolepsis as the combination of an embedded hanging topic and exceptional case marking

The previous section discussed prolepsis in terms of ellipsis. A major advantage of this analysis is that it solves the most pressing descriptive problem posed by this construction, namely the question how a verb can have two direct objects at once, a DP and a CP. The answer provided by the ellipsis analysis is that this is only apparently so: underlying, there are two verbs present (in separate clauses), one of which has been elided. The DP and the CP object each belong to one of these verbs, and hence there is no problem of double selection. However, as discussed, this also raises new difficulties. The most challenging one is that the semantics corresponding to the assumed structure does not convey the right meaning. In particular, the DP cannot always be interpreted as a thematic internal argument; rather it is perceived as an aboutness phrase with respect to the embedded clause. Another problem that is left unexplained is why the DP-argument has to be related to one of the arguments inside the embedded clause, and why that embedded argument has to be a pronoun.

This section aims to solve these issues by taking a different stance on the data. We argue that the proleptic DP is not an argument of the matrix verb at all, but related indirectly to the embedded verb. We claim that it is a left-peripheral topic doubled by a resumptive pronoun in the embedded clause. More precisely, we argue that prolepsis results from the interplay between embedded Hanging Topic Left Dislocation (HTLD) on the one hand and a form of Exceptional Case Marking (ECM) on the other. The proleptic ‘object’ is then a hanging topic exceptionally assigned accusative case by the matrix verb. It is coreferential with a resumptive pronoun within the embedded clause. The resumptive pronoun is the actual argument of the embedded verb.

This section is organized as follows. We start by introducing the required background for the analysis, namely the more general properties of HTLD. Section 4.2 provides a detailed analysis of the Middle Dutch prolepsis data. Section 4.3 compares the pattern from Middle Dutch to modern German and Dutch.

4.1. A few notes about Hanging Topic Left-Dislocation

HTLD, not to be confused with Clitic Left-Dislocation (CLLD) or Contrastive Left-Dislocation (CLD) is familiar from various languages, including Italian, Greek, and English. There is a prosodically isolated topic in the left periphery of the clause that is resumed by a pronoun within the clause. A simple example in English is (38):

(38) The angel, I heard him, speak.

The characteristics of HTLD have been described extensively in the literature (see among others Van Riemsdijk 1997, Grohmann 2003, Shaer & Frey 2004, Benincà & Poletto 2004, Alexiadou 2006, De Vries 2009). We will briefly review the most relevant ones here.

HTs are pragmatically aboutness topics. They always have a resumptive pronoun in the clause that agrees with it in φ-features but not (necessarily) in case. This becomes clear if we look at the examples in (39) and (40) from German and Icelandic, respectively. The HT appears in the default nominative case (see also Schütze 2001), whereas the resumptive pronoun displays the case that is in accordance with its syntactic function in the clause.

---

6 This is also noted by Massam (1985:181), who compares proleptic objects to of-NPs in English:

(i) I read of Carroll that she was painfully shy.
(39) Der Hans, ich kenne ihn schon seit zwölf Jahren.

‘As for Hans, I’ve known him for twelve years.’

(German, adapted from Van Riemsdijk 1997: 5)

(40) Pessi hringur, Ölafur hefur lofað Maríu honum.

‘This ring, Olaf has promised it to Maria.’

(Icelandic, Van Haaften, Smits & Vat 1983:135)

Secondly, the HT is base-generated in the outer left-periphery and not derived by movement. Boeckx & Grohmann (2005), for instance, show that German HTLD is not sensitive to islands. This is illustrated in (41): the HT der schöne Mann ‘the handsome man’ is separated from the resumptive pronoun ihn ‘him’ by a complex NP island.

(41) Der schöne Mann, Martin haßt [die Tatsache, dass die Frau ihn geküßt hat].

‘The handsome man, Martin hates the fact that the woman kissed him.’

(German, Boeckx & Grohmann 2005:141)

De Vries illustrates another argument for base generation on the basis of Dutch, namely that a HT cannot reconstruct into the base position of the resumptive pronoun (see also Grohmann 2003 for a discussion of German data). This becomes clear from the examples in (42) and (43). In (42) the reflexive pronoun zichzelf ‘himself’ in the left periphery cannot be bound by the subject of the clause Joop; in (43) a potentially expected Principle C effect does not occur, and coreference between the subject ze ‘she’ and the possessor Mieke contained within the object-related HT is fine. These judgments contrast sharply with those for the corresponding examples with CLD.

(42) *Pikante verhalen over zichzelf, Joop, hoort ze niet graag.

intended: ‘[As for] juicy stories about himself, Joop, doesn’t like to hear them.’

(Dutch, adapted from De Vries 2009:313)

(43) (Wat betreft) Miekes schoonvader, ik geloof dat ze, hem zelf niet zo mag.

as concerns Mieke’s father-in-law I believe that she him herself not so likes

intended: ‘(As for) Mieke’s father-in-law, I think that she, doesn’t really like him herself.’

(Dutch, adapted from De Vries 2014: 353)

We can add to this the impossibility of variable binding by a quantified phrase:

(44) *Zijn zuster, niemand, wilde haar verraden.

intended: ‘(As for) his, sister, nobody, wanted to betray her.’

In (44), the possessor zijn ‘his’ cannot covary with the quantified subject niemand ‘nobody’. Of course an unbound reading would be fine.
Crucially, it is possible in some languages to also have embedded hanging topics. This is illustrated in Italian in (45), where the HT *questo libro* ‘this book’ is resumed by the pronoun *ne*. It is to be noticed that the HT precedes the complementizer *che* ‘that’.

(45) Sono certa, [*questo libro, che non ne abbia mai parlato nessuno*].
I am certain this book that not of it has ever spoken nobody
‘I am sure that nobody has ever talked about this book.’

(Italian, Benincà & Poletto 2004:65, markings ours)

Considering the above facts as well as the prosodic isolation of hanging topics, we must conclude that they occupy a position outside the regular clause. In line with Benincà (2001) and Benincà & Poletto (2004) we therefore assume that a HT is generated in a discourse-related projection (a ‘high CP’ shell) above the regular CP-domain (say, Rizzi’s 1997 ForceP). For reasons that will become clear shortly, we do not assign a true parenthetical status to embedded HTs.

4.2. Prolepsis in Middle Dutch as HTLD plus ECM

With this information about HTLD in mind, reconsider a regular example of prolepsis, such as (46a). Presupposing that embedded HTLD is possible in Middle Dutch (we return to this below), we now analyze the sentence as in (46b), omitting irrelevant details.

(46) a. Si ontrieden den swerten here dat hi-ne soude te doet slaen.
they feared theACC blackACC lord that he-him would to death beat
lit. ‘They feared the black lord, that he would kill him.’

(Ferguut, v.1724-1725, translation ours)

b. [main.clause Si ontrieden [CP-high [DP den swerten here] C_high [CP-low dat hine soude te doet slaen]]]

The embedded clause is selected by the matrix verb *ontrieden* ‘feared’. The proleptic object is a hanging topic of the lower clause, base-generated in the higher CP shell. It is coreferent with the resumptive pronoun, which is the real external argument of the embedded main verb. Comparable to the situation in other ECM configurations, e.g. (47) in English, the HT (‘the black lord’) is ‘exceptionally’ assigned accusative case by the matrix verb.

(47) I asked [himACC to go].

---

7 More concretely, Benincà (2001:62) proposes the following cartography of the left periphery:

\[
\text{[ Disc } \text{ Hanging Topic [ Disc } \text{ che } [ \text{ Forcep excl wh- } [ \text{ Force } \text{ che } [ \text{ Topp (CLLD) Top [ Topp (che) [ Focp wh-Foc [ Foc } \text{ } [ Fupp [ Fup (che) TP ] ] ] ] ] ] ] ] ]]
\]

Remaining theory-neutral about the ‘cartographic program’ as such, we will simply assume that generalizations along these lines are at least descriptively correct.

8 The situation might be somewhat more complex than this. Koster (1999) argues that object clauses are in fact specifications of empty nominal objects, right-attached to a lower projection of the spine of the clause; see also De Vries (2010) for relevant discussion. However, such issues are tangential to the purposes of this paper, and we will represent embedded CPs as (right-hand) complements of the lower V, as is customary.

9 ECM can be described as follows: if the complement of a case-assigning predicate is non-nominal, the first nominal argument within the complement will be targeted (subject to locality constraints). In what follows, we will not be concerned with the exact syntactic mechanism of case marking. We may assume that it involves feature valuation under c-command as the result of an AGREE operation, but nothing crucially hinges on this.
A relevant difference between (46) and (47) is that the embedded clause is finite in the former, but non-finite in the latter. We would expect that case assignment to all arguments should be possible within the embedded clause in (46). This is correct, and, crucially, it is the resumptive pronoun that receives nominative, being the subject. Being dislocated, the HT is not syntactically an argument of the predicate itself; hence, it does not receive case within the embedded clause. Since it is located in the highest projection, and hence in the ‘edge’ of the phase/cycle, it is visible from above. Therefore it is quite plausible that it can be accusative-marked by the matrix verb, without any conflict in case features. Thus, because of the HT configuration, the proleptic accusative can be analyzed as involving ‘exceptional ECM’.

Let us spell out this last point in a little more detail. We claim that a hanging topic of an embedded clause can be assigned case by the matrix verb. Since case marking is usually done within one and the same phase, this might seem strange at first glance, but it is not upon closer scrutiny. The exceptional instance of ECM is made possible by a number of interacting factors. Recall from the discussion in section 4.1 above that HTs normally appear in the default case. In line with Schütze (2001), we could take this to mean that they have not been assigned case syntactically, and simply receive the default value for case in the morphological component as a last-resort strategy. However, for embedded HTs the situation is different: for these, the default strategy does not need to be invoked, but the starting point is the same. Namely, they are not case-marked within their own clause because they are not directly selected for by a case-marking predicate. Being generated in the leftmost specifier of the embedded clause, i.e., the edge of the embedded C°-phase, they are not yet sent off to spell-out (and hence to the morphological component) when the matrix v° is merged. For practical purposes, the edge of a phase can be considered to be part of the next-higher phase – in this case, that of the matrix v°. Consequently, the HT may be assigned case by the matrix verb in principle: it is caseless upon merger of the verb and it is in the right domain (locally c-commanded by the relevant verbal head). A prerequisite for this is that the verb is of the right lexical class, i.e., the class of ECM-verbs.

In some cases the situation is slightly more complex due to movement of the proleptic object. Consider the repeated example in (48), which we now analyze as in (49):

(48) *Den* †*inghel* horden wi nochtan *hoe hi* sprac den wiven an.  
the_{ACC} †angel heard we still how he spoke the women to  
*lit.* ‘The angel, we still heard how he talked to the women.’  
(Spieghel Historiae I: VII - XXXVI, v.17-18, translation ours)

(49) [main.clause [DP *Den* †*inghel* horden wi nochtan horden [CP-high *den* †*inghel* C_{high} [CP-low *hoe C_{low} [TP hi sprac den wiven an]]]]

---

10 It is worth noting that Polinsky & Potsdam discuss a related phenomenon in Tsez. Consider the examples in (ia/b), from Polinsky & Potsdam (2001:584).

(i) a. enir užā magalu bāc’ruli r-iyxoy  
mother [boy-ERG bread.III.ABS III-ate-NMLZ].IV IV-knows  
‘The mother knows the boy ate the bread.’

b. enir užā magalu bāc’ruli b-iyxoy  
mother [boy-ERG bread.III.ABS III-ate-NMLZ] III-knows  
‘The mother knows the boy ate the bread.’

In Tsez, transitive main verbs agree with their objects. In (ia), iyxoy ‘know’ agrees with the object clause užā magalu bāc’ruli ‘the boy ate the bread’, indicated by the prefix r- for class IV. In (ib), on the other hand, the verb agrees with an argument inside the embedded clause, namely the object magalu ‘bread’, indicated by the prefix d- for class III. Polinsky & Potsdam (2001:609f) convincingly argue that agreement between the main verb and the embedded object is only possible, and in fact obligatory, if the embedded object is the topic of its own clause (and either overtly or covertly moved to a peripheral topic position).
Again, the proleptic object, here *den inghel* ‘the angel’, is a HT base-generated in the left periphery of the embedded clause. In this position it gets assigned accusative case by the matrix ECM-verb *horden*. The subject pronoun *hi* ‘he’ within the embedded clause is a resumptive pronoun, co-indexed with *den inghel*. What is different from (46b) is that the proleptic object has undergone further topicalization into the left periphery of the matrix clause in this particular sentence. Recall that in Dutch, basically any constituent of the clause can be topicalized for reasons of highlighting or discourse linking. This linearly first constituent is then followed by the finite verb in the second position (V2), resulting in linear inversion with the subject, which is in SpecTP.

Even though the HT is part of an embedded finite clause, it is important to see that it is in principle available for further movement simply because it is in the highest specifier, hence in the edge of its phase. We have encountered not only topicalization into the matrix, but also examples in which the proleptic object surfaces in the middle field of the matrix. All of these involve positions in which regular objects can appear as well, cf. (10) or (15).

The analysis provides a natural explanation for several properties of the prolepsis construction discussed above. First of all, recall that the proleptic object seems to act as an aboutness topic. The sentence in (46), for instance, presumably means that the subjects feared something about the black lord (rather than the black lord himself), namely that he would kill someone. This is exactly the meaning that is expected under the present analysis, since HTs are aboutness topics (cf. the references in section 4.1).

Secondly, from the assumption that we are dealing with HTLD in these sentences it also follows that the relevant argument in the embedded clause is necessarily a pronoun and that this pronoun is coreferent with the proleptic object.

Thirdly, we have seen many examples where the proleptic DP agrees in φ-features with the resumptive pronoun, but not in case. This is expected if the proleptic DP is a HT: a HT acquires case by another means than the related resumptive pronoun (as discussed above), and they are not part of the same movement chain.

Fourthly, a HT may be in a nonlocal configuration with respect to the corresponding resumptive pronoun, i.e., the pronoun can be contained within an island; recall (41), for instance. This is also seen in several of the Middle Dutch examples of resumptive prolepsis, as well as in Greek (17). Consider (48) again from this perspective. In this example, the proleptic DP *den inghel* ‘the angel’ and the resumptive pronoun *hi* ‘he’ are separated from each other by a *wh*-island boundary (note the question word *hoe* ‘how’). The same phenomenon can be observed in the example in (15), not repeated here.

Fifthly, from the assumption that the proleptic object gets accusative case via an exceptional case of ECM, it follows that (i) the class of verbs that appears in this construction overlaps with the class of ECM-verbs, and (ii) that the case on the proleptic object is normally accusative (but see the next subsection for a principled exception in passive constructions).

In short, we showed that the current analysis accounts for both the syntactic and semantic properties of prolepsis without introducing any theoretical stipulations.

4.3. A comparative Germanic view on prolepsis

This final subsection touches upon some comparative issues of the prolepsis construction. First, it is remarkable that the proleptic accusative construction is not possible in modern Dutch. Consider the examples in (50a/b), which directly correspond to (46a) and (48) in Middle Dutch. Neither is acceptable – contrary to periphrastic solutions cited in (8).
(50) a. * Zij vreesden de zwarte heer, dat hij hem zou doden.  
they feared the black lord that he him would kill  
intended: ‘They feared [of] the black lord, that he would kill him.’  

b. * De engel hoorden wij hoe hij de vrouwen toesprak.  
the angel heard we how he the women spoke to
intended: ‘[As for] the angel, we heard how he talked to the women.’

This can be explained straightforwardly, since modern Dutch does not allow embedded dislocation or even topicalization in general. Consider (51a/b), for instance:

(51) a. * Peter zag (in de tuin), dat (in de tuin), de meisjes t/er, zaten.  
Peter saw in the garden that in the garden the girls sat  
intended: ‘Peter saw that in the garden, the girls were sitting (there).’  

b. * Peter zag (de meisjes), dat (de meisjes), Joop t/ze, aansprak.  
Peter saw the girls that the girls Joop addressed  
intended: ‘Peter saw that the girls, Joop addressed (them).’

Whether there is a resumptive pronoun or an unpronounced trace/copy, and whether the displaced constituent is put before or after the complementizer, all examples are unacceptable.11 Why this is so is an issue that is orthogonal to the purposes of this paper (see Zwart 1997:245ff for a partial explanation). What is relevant is that such possibilities are a prerequisite for the proleptic construction under the proposed analysis. If these requirements are not fulfilled, it is correctly predicted that prolepsis is impossible.

Returning to Middle Dutch, we expect it to be more liberal than modern Dutch with respect to embedded displacement. This is indeed what we find. A relevant illustration of embedded HTLD is cited in (52). Here, the HT in sine nueseugen ‘in his nostrils’ is related to the pronoun re ‘there’ in the embedded clause.

11 Note that the illustrations in the main text are not to be confused with the phenomenon of focus scrambling, which is allowed:

(i) ... dat ZO’N AUTO zelfs DE JONGENS niet mooi vinden.
that such a car even the boys not nice find:PL  
‘...that not even the boys like such a car.’

According to Neeleman (1994), Zwart (1997), and others, scrambling of this type – or in fact any type, we would say – does not involve embedded topicalization into the left periphery; rather, the object is moved up in the middle field, potentially leaving the subject behind in a lower position.

There is one more complication of (colloquial) Dutch we will ignore here, which is the possibility of restarting a sentence ‘below’ the complementizer, which leads to a kind of apparently embedded main clause phenomena. In particular, it is possible to construe cases of contrastive left-dislocation for some speakers; see (ii), taken from Zwart (1997:251). Here, Marie is the preposed object of the lower clause.

(ii) Jan zei dat Marie (die) kuste hij niet.  
John said that Mary that-one kissed he not  
‘John said that Mary, he did not kiss.’

Within the ‘embedded’ domain after the complementizer, the finite verb kuste ‘kissed’ is in the second position, as if it concerns a main clause. If the verb is positioned final, as in regular embedded clauses, the example becomes completely unacceptable. This is expected, considering (50) in the main text. From the perspective of our analysis, (ii) is irrelevant for various reasons. Most importantly, whatever is going on in these cases, is still below the complementizer, whereas in the proleptic construction, material is raised above C.
(52) Men pleghet in sine nuesegeten dat me-re eenen ring in doet.  
     one is.used.to in his nostrils that one-there a ring in does  
     lit. ‘One is used to the fact that in his nostrils, one puts a ring therein.’  
     (Stoett 1923:246, translation ours)

The example provides an instance of ‘regular’ embedded left-dislocation rather than prolepsis proper, since the fronted phrase is preceded by a preposition, which is a case assigner itself (see also Benincà & Poletto 2004).

Interestingly, we find a similar phenomenon in variants of modern German (which is also more liberal than modern Dutch with respect to scrambling across the subject). Example (53) illustrates embedded left-dislocation according to Grohmann (1997):

(53) Der Bauer glaubt, diesen Frosch, daß sie den gestern geküßt hat.  
     the farmer believes this frog, that she it yesterday kissed has  
     lit. ‘The farmer believes that this frog, she kissed it yesterday.’  
     (adapted from Grohmann 1997:13)

In fact, (53) may already be an instance of prolepsis. Note that we found examples in Middle Dutch where the resumptive pronoun has an object role rather than a subject role; see (12), for instance. Moreover, and this is crucial, a regular hanging topic would have nominative case; cf. (39). In (53), however, the HT is accusative-marked, as in the prolepsis construction.

In order to confirm if the proleptic accusative really exists in modern German, we performed a small survey. The outcome confirms our first impression based on (53) – again, for a subset of speakers. Consider the example in (54), which is similar to (48) in Middle Dutch:

(54) Den Engel hörte ich wie er zu der Frau sprach.  
     the Angel heard I how he to the woman spoke  
     lit. ‘I heard the angel, how he spoke to the woman.’

Here, the topicalized accusative phrase den Engel ‘the angel’, is related to the resumptive subject er ‘he’ in the embedded clause. As before, we propose that the proleptic object is generated as a hanging topic of the embedded clause, and receives accusative case from the matrix ECM-verb in this position (prior to further topicalization into the left periphery of the matrix).

Interestingly, German provides us with the opportunity to check whether the accusative case of the proleptic object is indeed assigned by the matrix verb. To this end, we can use passivization as a test. We predict that if the matrix clause is passivized, the proleptic object will be promoted to the subject position, and consequently receives nominative case. Although we have not yet discovered such examples in Middle Dutch, this is indeed what we find in modern German; see (55):

\[
(55) \text{Der Bauer glaubt, diese} \text{nen} \text{Frosch, daß er den gestern} \text{geküßt} \text{hat.}  
\]

\[
\text{the farmer believes this frog, that he} \text{it yesterday} \text{kissed has}  
\]

\[
\text{lit. ‘The farmer believes that this frog, she kissed it yesterday.’}  
\]

\[\text{adapted from Grohmann 1997:13}\]

---

12 According to Grohmann (1997), this is possible in those dialects of German that also allow extraction out of embedded daß-clauses.

13 Four out of seven speakers we consulted accepted the relevant sentences to some degree. We will leave a more comprehensive description of the distribution of the proleptic accusative construction in the Germanic language area as a topic for further research.
(55) a. *Der Engel würde gehört wie er zu der Frau sprach.
It was heard how the angel spoke to the woman.’

b. *Den Engel würde gehört wie er zu der Frau sprach.

In (55a), der Engel ‘the angel’ has nominative case; (55b) shows that an accusative is not acceptable in this passive configuration This implies that the accusative cannot be viewed as a default case in the proleptic construction, and hence that its presence in (53) and (54) must be due to ECM marking.

To summarize, this section has provided an explanation of why modern Dutch lacks the prolepsis construction attested in Middle Dutch, given the analysis proposed in section 4.2. Furthermore, we showed that the proleptic accusative is still found acceptable in (varieties of) modern German. Although more cross-linguistic research remains to be done, our initial investigation provided empirical results that clearly confirm the proposed analysis.

5. Conclusion

We presented an empirical and theoretical analysis of the so-called proleptic accusative construction. This puzzling phenomenon involves two constituents, a noun phrase and a clause, that apparently compete for the object status in the matrix clause. The relevant noun phrase, or ‘proleptic object’, appears to be accusative-marked by the matrix verb, but it is thematically related to the embedded clause, and coreferent with a resumptive pronoun within that clause. Often, the syntactic function of the resumptive pronoun (possibly pro, depending on the agreement system in the particular language) is subject, but it can also be object. The proleptic constituent is always marked with accusative case (possibly abstractly, depending on the morphological properties of the particular language). But its location in the sentence is variable. Two common positions are the one directly preceding the embedded clause, or the one topicalized within the main clause. As for the type of embedded clause, it can be a regular declarative complement introduced by that, but it can also be an embedded question introduced by a wh-constituent.

We discussed the proleptic accusative in Middle Dutch in some detail, making use of a small corpus of examples established by Stoett (1923). Thereby, we expanded the knowledge of the properties of the construction type, building on a small body of literature that mainly addresses ancient Greek and Latin, and Nahuatl. Based on empirical descriptions so far, the impression one gets is that prolepsis is rare – both in terms of cross-linguistic distribution and frequency. This may well be related to the fact that its derivation requires an uncommon combination of syntactic mechanisms. But it is probably also the case that prolepsis is a slumbering possibility in a number of languages, and as such empirically underexposed. Based on a small-scale survey, we found it to exist in modern German, whereas reference to this is lacking in grammars or other linguistic literature, as far as we are aware. Contrary to the situation in (varieties) of modern German, the phenomenon is absent in modern Dutch. We related this to the loss of embedded topicalization and dislocation, which was still possible in Middle Dutch.

Theoretically, we argued that the proleptic accusative cannot be explained in terms of (copy) raising for a variety of reasons. To rehearse just one problem, the alleged base-position of the proleptic object, where the resumptive pronoun resides, can be inside a (wh-)island, which makes movement highly unlikely. Furthermore, we showed an example from Middle Dutch involving a coordinated object and split coreference via two separate resumptive pronouns with differing syntactic functions inside the embedded clause. Such
cases cannot easily be explained in terms of a movement chain. As an alternative, we worked out a novel analysis involving clausal coordination and ellipsis, such that the ‘competing objects’, DP and CP, are distributed over the two separate matrix clauses, each selected by an instance of the selecting matrix verb. Although this straightforwardly solves the syntactic complications discussed, it leads to problems with semantic interpretation in various cases. The core issue here is that the proleptic constituent is usually not interpreted thematically as an internal argument of the matrix verb, \(^{14}\) but rather as an aboutness phrase with respect to the embedded clause.

Taking seriously the aboutness meaning aspect of the proleptic accusative, which is also reflected explicitly in periphrastic variants thereof, we developed a final analysis in section 4 of the paper. \(^{15}\) We link aboutness to hanging topic left-dislocation being involved, considering, among other things, that hanging topics are always aboutness topics. What is somewhat unusual from a cross-linguistic perspective is that HTLD then targets the embedded clause. There are however good reasons to assume that it is correct, nevertheless. It is also to be noted that the type of matrix verbs involved (often verbs of perception and cognition) facilitates pragmatic plausibility for a discourse in which an entity within the embedded environment figures as a sentence topic. Moreover, we showed that the syntactic possibility of embedded topicalization and dislocation in a particular language is a necessary prerequisite for prolepsis.

Thus, the proleptic object, i.e., the seemingly additional accusative argument in the matrix, is base-generated as an embedded hanging topic (HT) of the complement clause involved. We argued that this has a series of welcome consequences. First, it explains the presence of a resumptive pronoun in the embedded clause. This pronoun necessarily occupies an argument position within that clause. In principle, it can be any argument, but of course there is a clear preference for subjects to be topical. Secondly, the relationship between the proleptic HT and the pronoun is one of coreference, not movement. Therefore no locality constraints are expected to play a role. Again, this is in line with empirical findings. Furthermore, the HT and the pronoun may differ in case. Thirdly, since the embedded HT is in the edge of the lower clause, it is expected to be available for further movement into the matrix. Indeed, various such cases are attested. As already commemorated, a relatively frequent configuration is one where the proleptic HT is further topicalized into the left periphery of the matrix. Fourthly, the embedded HT must receive case via exceptional case marking. We argued that this is possible in principle because there is no phase boundary between the matrix verb and the topic high up in the left edge of the lower CP. Consequently, the case marking on the proleptic HT will be accusative. If we are correct, this implies that the phenomenon of ECM is more general than thought so far, an interesting conclusion that seems to open up future directions of research. Apart from that, we showed that there is one exception to marking with accusative. Namely, if the matrix is passivized, promotion to subject is triggered, and consequently nominative case marking.

---

\(^{14}\) There is no reason to assume that the ellipsis analysis is excluded across-the-board, as it might coexist with the – generally preferred – alternative in terms of HTLD for those examples (and only those) in which it reflects a semantically adequate structure. The grammar, as an automatic system, does not choose between analyses, it only provides possibilities, and individual examples can in principle be structurally ambiguous, with corresponding meanings.

\(^{15}\) At first sight, we run into a somewhat paradoxical situation. The analysis with the most straightforward syntactic solution leads to semantic complications, but the semantically correct alternative requires a rather complex syntax. On closer inspection, however, the situation is far from symmetrical, as explained in the main text: the required complex syntax in the latter option is actually well-founded, but the semantic complications in the former are a serious drawback that cannot easily be explained away.
To conclude briefly, we argued that prolepsis is the interesting result of interacting syntactic processes, namely embedded HTLD combined with an exceptional form of ECM. Importantly, there is no need to invoke construction-specific stipulations: as one would expect, the properties of the proleptic accusative can be derived from more basic ingredients of the grammar.

References


**Literary works included in the corpus**

<table>
<thead>
<tr>
<th>Title</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Der minnen loep</em></td>
<td></td>
</tr>
<tr>
<td><em>Ferguut</em></td>
<td></td>
</tr>
<tr>
<td><em>Leven van Sinte Amand II</em></td>
<td></td>
</tr>
<tr>
<td><em>Rijmbijbel</em></td>
<td></td>
</tr>
<tr>
<td><em>Roman van Walewein</em></td>
<td></td>
</tr>
<tr>
<td><em>Sinte Franciscus Leven</em></td>
<td></td>
</tr>
<tr>
<td><em>Spieghel Historiae I &amp; III</em></td>
<td></td>
</tr>
<tr>
<td><em>Vanden Gheesteliken Tabernakel</em></td>
<td></td>
</tr>
</tbody>
</table>

24