Subject Relatives and Expletives in Early New High German

CAITLIN LIGHT
University of Pennsylvania

Introduction

The status of the subject position in German has been the source of some debate.\textsuperscript{1} For example, some studies (Biberauer 2004; Richards and Biberauer 2005) have argued that German does not have an EPP requirement in the traditional sense. The absence of an expletive that occurs specifically in the subject position (as opposed to the topic position in Spec,CP) seems to support the argument that Spec,TP has no special status in German.

This paper will argue against such analyses, and show that in historical stages of German, we see evidence of a subject expletive licensed specifically to fill Spec,TP. This expletive, \textit{da} in Early New High German (ENHG), is merged specifically when the logical subject does not move to Spec,TP, leaving the position empty. This supports a traditional analysis of the EPP in German. Furthermore, I will show that the existence of expletive \textit{da} lends support to the argument that two (non-topic) subject positions are available in the German clause structure (cf. Haeberli 1999, 2000, 2005), which I take to be Spec,TP and Spec,vP (the base position of the subject).

This study is based on data from a parsed corpus of Martin Luther’s \textit{Septembertestament}, a translation of the New Testament published in 1522. Luther intended for his Bible translation to be accessible to a wide audience, and hence the text represents a more colloquial sample of ENHG. The \textit{Septembertestament} corpus, at the time of this study, consisted of approximately 40,000 words that have been fully POS-tagged and parsed. This provides a sample of 1,716 subordinate and 2,996 main clauses.

The paper will proceed in the following way. In Section 1, I will outline the behavior of \textit{da} in ENHG. Section 2 will introduce some related constructions else-

\textsuperscript{1} I would like to thank Anthony Kroch, Julie Legate, Dave Embick, and Joel Wallenberg for their assistance in this work. I am also deeply grateful to Beatrice Santorini and Florian Schwarz, who provided the Modern German judgments used in this paper. Of course, any errors are my own.
The behavior of da in Early New High German

The pronominal adverb da is generally locative, although it can also have a temporal interpretation, dependent on context. In both ENHG and contemporary German, da is also available as a complementizer. The following examples show ENHG da behaving as a locative (1) and temporal (2) adverb, and a complementizer (3).

1. Vnd es waren da viel weyber . . .
   and it was DA many women . . .
   ‘And there were many women there.’
   (Septembertestament, Matthes 27:55)

2. Da berieff Herodes die weysen heymlich . . .
   DA appointed Herod the wise secretly
   ‘Then Herod secretly appointed the wise men.’
   (Septembertestament, Matthes 2:7)

3. vnd da sie yhn sahen, fielen sie fur yhn nyder . . .
   and DA they him saw fell they before him down
   ‘And when they saw him, they fell down before him.’
   (Septembertestament, Matthes 28:17)

Note that, like other demonstratives (such as the personal demonstrative pronouns der, die, and das), locative da may also be used as a relative pronoun in ENHG, alternating with the wh-word wo ‘where.’ I do not count this as a separate property of da because it behaves like other demonstratives in this sense.

All of these facts are essentially unchanged between ENHG and contemporary German. However, the ENHG element da differs in one significant respect: in subordinate clauses, it shows a high correlation with the position of extracted subjects, particularly in relative clauses and free relatives.

4. Simon, der do heyst Petrus
   Simon who DA is-called Peter
   ‘Simon, who is called Peter’
   (Septembertestament, Matthes 4:18)

5. Selig sind, die da geystlich arm sind
   blessed are who DA spiritual poor are
   ‘Blessed are those who are spiritually poor.’
   (Septembertestament, Matthes 5:3)
The use of *da* in these examples sounds archaic, and in fact distinctly bibli-
cal, to native German speakers. They are not acceptable as contemporary German
utterances, even if they are adjusted to exclude DP extraposition and other construc-
tions which are no longer grammatical in German. The behavior of *da* has clearly
undergone some change since the ENHG period.

In subordinate clauses, *da* is almost entirely restricted to clauses with gaps.
Even more significantly, only subject gap clauses show this effect. Out of the 79
subordinate clauses where *da* occurs, only three appear in clauses without a subject
gap.

Figure 1: Appearance of *da* in different types of subordinate clauses

<table>
<thead>
<tr>
<th></th>
<th>da</th>
<th>No da</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clauses with gaps</td>
<td>77</td>
<td>422</td>
<td>499</td>
</tr>
<tr>
<td>All other subordinate</td>
<td>2</td>
<td>1146</td>
<td>1148</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>79</td>
<td>1568</td>
<td>1647</td>
</tr>
</tbody>
</table>

Figure 2: Appearance of *da* in clauses with gaps

<table>
<thead>
<tr>
<th></th>
<th>da</th>
<th>No da</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject gap</td>
<td>76</td>
<td>326</td>
<td>402</td>
</tr>
<tr>
<td>Non-subject gap</td>
<td>1</td>
<td>96</td>
<td>97</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>77</td>
<td>421</td>
<td>499</td>
</tr>
</tbody>
</table>

In the subject gap examples where *da* appears, it is uniformly clause-initial. In
contrast, the single example of *da* in a clause with a non-subject gap is not clause-
initial, shown in (6). This example is a fairly unremarkable example of *da* as a
locative adverb. The behavior of *da* in subject gap clauses seems to suggest an
entirely different phenomenon.

(6) *vnd wie viel korbe hubt yhr da auff?*
    and how many baskets lifted you DA up
    ‘And how many baskets did you collect there?’
    (*Septembertestament*, Matthes 16:9, 16:10)

Although *da* is available both as a complementizer and a relative pronoun in
ENHG, as I discussed above, I will argue that the *da* correlated with subject gaps
is neither. This *da* almost always occurs below an overt relative pronoun, making
it clear that it cannot itself be a relative pronoun. In addition, when it is used as a
relative pronoun (as discussed briefly above), *da* has a locative interpretation, and
corresponds to the extraction of a locative adjunct, not a subject.

Further evidence demonstrates that this *da* is not a complementizer. The corpus
provides two examples of “subject gap” *da* occurring below an overt complement-
izer, as shown below.

(7) a. vnd wirtt eyn bawm das *da* komen die vogel vnder dem hymel
and becomes a tree that DA come the birds under the heaven
‘And (it) becomes a tree, so that the birds of the heavens come.’

(Septembertestament, Matthes 13:32)

b. Wer sagen die leutt, das *da* sey des menschen son?
who say the people that DA is of-the man son
‘Who do the people say is the son of man?’

(Septembertestament, Matthes 16:13)

In (7a), an adverbial clause headed by the complementizer *das* (Modern German
*daß*) has clause-initial *da* in the subject gap created by extraposition of the subject
(evident because the subject DP occurs after the finite verb). In (7b), the subject of
a complement clause is extracted, and *do* (an alternate spelling of *da*) appears in
the clause of its origin, again below the complementizer *das*. Based on these two
examples, we can safely exclude the possibility that the *da* correlated with subject
gaps is a complementizer. The evidence calls for an alternative analysis.

2 Subject gaps in Germanic

This phenomenon is not wholly unique cross-linguistically. Several languages in
the Germanic family have constructions in which subject gaps left by extracted
subjects are filled by expletive elements. In order to provide a context for the puzzle
at hand, I will review two such cases. A better understanding of the mechanisms
underlying these phenomena may help explain the behavior of *da* in ENHG.

2.1 Danish *der*

In Danish, the element *der* appears clause-initially in relative clauses and indirect
questions with a subject gap (cf. Jacobsen and Jensen 1982; Erteschik-Shir 1984;
Vikner 1991; Mikkelsen 2002). The following examples demonstrate that *der* is
grammatical in an indirect question where the subject has been extracted, but not
permitted in a similar example if the object has been extracted instead.

(8) Jeg ved ikke hvem *der* kan li’ ham.
I know not who DER likes him
‘I don’t know who likes him.’
(9) * Jeg ved ikke hvem der han kan li’.
    I know not who DER he likes
    ‘I don’t know who he likes.’
    (Erteschik-Shir 1984)

The literature is divided on whether this element is the subject expletive *der* ‘there’ occupying the gap (Erteschik-Shir 1984; Mikkelsen 2002), or a homophonous element in C (Jacobsen and Jensen 1982; Vikner 1991). There are two obvious reasons to position *der* in the subject gap (Spec,TP). The first is that *der* does not appear in clauses with a non-subject gap. The second is that *der* is independently attested as a subject expletive in Danish, so it is clearly available to occupy Spec,TP.

Mikkelsen (2002) shows that the distribution of *der* supports this analysis. It is marked, but acceptable, for *der* to occur with the complementizer *at* and the element *som*, which Mikkelsen analyzes as an invariant operator (like OP) in Spec,CP. Whenever any of these elements co-occur, they *must* be in the order *som at der*, requiring *der* to appear below the complementizer. The analysis of *der* as a C-element requires CP-recursion to account for this structure, while idiosyncratic properties of each head must explain the ordering restriction. However, the assumption that *der* is in Spec,TP predicts exactly this distribution.

Vikner (1991) suggests some problems with the expletive analysis which must be addressed. There is generally a transitivity restriction on the appearance of the expletive, as well as a restriction on the definiteness of the logical subject. When *der* appears in the subject gap position, it defies both of these restrictions.

(10) a. vi kender de lingvister der vil læse denne bog
    we know the linguists DER will read this book
    ‘We know the linguists who will read this book.’
    (Vikner 1991)

Because German (unlike Danish) does have transitive expletive constructions, the first issue is unique to the Danish case, and a discussion of it is beyond the scope of this paper. I will focus instead on the definiteness problem. In response to Vikner’s point, Mikkelsen proposes that the gap in relative clauses behaves like an indefinite, even when its antecedent is not. This has been previously argued, particularly for English (cf. Browning 1987; Bianchi 1999), based on examples like the following:

(11) a. * There were the men in the garden

    b. The men that there were in the garden were all diplomats
    (Browning 1987)

I find the judgment in (11b) tricky, but agree that it is certainly more acceptable than (11a). Furthermore, in isolation, the phrase *the men that there were* seems
relatively natural. Mikkelsen (2002) proceeds by relating this to the argument made in Reinhart (1987), that relative *wh*-words in English are inherently indefinite. Mikkelsen suggests that the evidence provided by *der* motivates us to extend this analysis to Danish, to explain why the relative pronouns seem to be patterning as indefinites.

2.2 Yiddish subject gaps

Yiddish is a symmetric V2 language: verb-second orders are required both in matrix and subordinate clauses. If the subject is not raised to the topic position, the expletive element *es* may appear instead. However, if another constituent is topicalized, *es* is not licensed. More importantly for our purposes, the expletive also appears in subject gap position, when the subject is extracted from a free relative or indirect question (cf. Diesing 1990, 1997; Prince 1993, 1989). This does not take place in ‘ordinary’ relative clauses.

(12) a. 
   Ikh veys nit ver *es* iz gekumen  
   I know not who ES is come  
   ‘I don’t know who came.’

b. * Ikh veys nit ver iz gekumen

c. Ikh veys nit vos Max hot gegesn  
   I know not what Max has eaten  
   ‘I don’t know what Max ate.’

d. * Ikh veys nit vos Max hot *es* gegesn  
   (Diesing 1990)

(13) a. Der melamed vos iz besser far ir iz beser far mir.  
   the teacher that is better for her is better for me  
   ‘The teacher that is better for her, is better for me.’

b. * Der melamed vos *es* iz beser far ir iz beser far mir.  
   (Prince 1989)

Free relatives and indirect questions clearly have a special status here. Prince (1989) argues that the relevant characteristic of these clause types is that the extracted element represents brand-new information in the discourse, unlike relative clauses, which have a clear referent. This allows us to relate the expletive in subject extraction contexts to the expletive in the Yiddish ‘subject postposition’ construction. As in the subject gap examples, subject postposition triggers *es* insertion in the case that no other element is topicalized.

(14) a. *a yid* geyt farbay oyf-n gas.  
   a guy goes by on-the street  
   ‘A guy passes by on the street.’
Using a small corpus of narrative Yiddish, Prince (1989) demonstrates that subject postposition in Yiddish is more frequent with brand-new subjects. Prince argues that in indirect questions and free relatives, the extracted element is pragmatically treated as brand-new, because it is not related to any elements already present in the discourse. As a result, they are extracted from the postposed position. The insertion of the expletive *es* in these cases is required because there is no subject trace in the topic position, and the expletive is licensed to fill the gap.

This analysis cannot be directly applied to the ENHG data under consideration. Unlike *es* in Yiddish, ENHG *da* is attested in ordinary relative clauses. In addition, because Yiddish is a symmetric V2 language, we know that there is a requirement to fill the topic position in subordinate clauses. In German, subordinate clauses have no topic position. Expletive *es* is frequently ungrammatical in embedded contexts, where it would be required in a matrix clause, because it is generally licensed to fill the topic position in Spec,CP. This raises the question of whether the subject position must be filled in German subordinate clauses. The present study will, in fact, prove relevant to this issue.

### 3 Previous analyses of *da*

The previous section has shown several Germanic languages which require an expletive to fill a gap left by a subject. This seems to fit with the facts we have seen for ENHG. However, as I will demonstrate, *da* is clearly not an expletive in Modern German as it is spoken today.

In normal usage, *da* may carry either a locative or temporal interpretation according to context, much like *there* in English (cf. Bayer and Suchsland 1997; Koeneman and Neeleman 2001; Kratzer 2004; Richards and Biberauer 2005). This ambiguity of interpretation often makes it difficult to identify whether *da* may be semantically null in a clause. Despite this, I will provide evidence that *da* cannot be analyzed as an expletive in Modern German.

Bayer and Suchsland (1997); Richards and Biberauer (2005) have proposed that *da* may in fact be used as a subject expletive in Spec,TP, which may be merged into the subject position when the logical subject remains low.

(15) a. daß *(da)* gestern ein Schiff versunken ist
    that (DA) yesterday a ship sunk is
   ‘…that a ship sunk yesterday.’

b. *daß da ein Schiff gestern versunken ist

c. daß ein Schiff gestern versunken ist
   (Richards and Biberauer 2005)
In (15a), we see that the proposed expletive *da* may appear if an adverb (*gestern*) intervenes between it and the logical subject, which must therefore (presumably) be low. If the logical subject appears to the left of the adverb, as in (15b), the sentence becomes ungrammatical, while (15c) shows that it is perfectly acceptable when *da* is omitted. Richards and Biberauer (2005) argue that this is evidence that *da* is an expletive that may fill a high subject position. They argue that because the EPP in German does not behave as it does in languages like English, filling this position is optional; but if the subject raises from its base position, the expletive element should be disallowed.

This is an attractive analysis for *da*. However, it must ultimately be rejected, for contemporary German at least. With further exploration, it becomes evident that native German speakers easily accept the sentence in (15b), contrary to the judgments in Richards and Biberauer (2005). Furthermore, sentences such as these are not possible without a locative (or temporal) interpretation for *da*, which cannot be semantically null. When paired with another locative, the sentence generally gains a redundant or contradictory reading.

(16)  

(a)  

I believe that DA at the party danced became  
‘I believe that there was dancing there, at the party.’

(b)  

?? Ich glaube, daß auf der Party *da* getanzt wurde.

(17)  

Ich glaube, daß *da* in China angerufen wurde.  
I believe that DA in China up-called became.  
‘I believe that someone in China was called (from there).’

Although the examples in (16) are not ungrammatical, the presence of *da* is considered redundant, pointing to the same location as the adjunct *auf der Party*. In contrast, (17) is interpreted such that *da* and *in China* are actually indicating different locations; while China is identified as the place where the call was received, *da* is consistently interpreted as referring to the place where the call was made – that is, the location of the agent. These examples provide evidence that *da* cannot have the semantic properties of an expletive in contemporary German.

Furthermore, the structural position of *da* does not fit this analysis. In German, weak pronouns are necessarily high in the structure. If an object pronoun is also included in the sentence, it must occur to the left of *da*.

(18)  

(a)  

* Ich glaube, daß *da* er ihn gegessen hat.  
I believe that DA he it eaten has  
‘I believe that he ate it there.’

(b)  

* Ich glaube, daß er *da* ihn gegessen hat.

(c)  

Ich glaube, daß er ihn *da* gegessen hat.
The position of *da* thus does not match with the claim that it occupies Spec,TP. What we see instead is that in Standard German, *da* is weak, and behaves much like weak pronouns in German. As a result, it prefers a position at the left periphery of the clause, often string-adjacent to the material in C. However, it must appear after a pronominal object. Weak elements in German follow an ordering hierarchy: weak adverbs like *da*, for example, must be lower than a weak object pronoun (cf. Lenerz 1977). The fact that *da* tends to appear at the left periphery is completely expected, simply because it is weak.

An alternative analysis of *da* has been proposed by Kratzer in the literature on situational semantics, which captures the ambiguous interpretations associated with the element. She suggests that *da* may be considered a *situation pronoun*, an adverbial pro-form which may be used to refer to any salient information in the context situation.

(19) Was riecht denn *da* so komisch?
    what smells PART DA so funny
    ‘What’s the strange smell here?’

(20) *Da* brandelt was.
    DA burns something
    ‘Something is burning.’
    (Kratzer 2004)

As I have noted before, non-subject gap *da* in ENHG has the same general properties as *da* in Modern German, including an ambiguous and context-dependent interpretation. I will therefore assume that the general usage of *da*, both in ENHG and in Modern German, falls under Kratzer’s analysis (excluding, of course, its use as a complementizer or relative pronoun). However, as it stands, this analysis offers no further insight into the particular correlation between *da* and subject gaps. Instead, in the next section, I will show that the necessary evidence can be found farther in the past.

4 The historical origins of expletive *da*

Regular sound change led to the collapse of two OHG adverbs, *thō* ‘then,’ and *thār* ‘there,’ into a single adverb *da*. This led to the element’s ambiguous interpretation (cf. Axel 2007), allowing it to develop into a situation pronoun by the modern period. However, only the behavior of OHG *thār* offers insight into the curious behavior of subject gap *da* in ENHG. In fact, *thār* shows a correlation with subject gaps, just as we have found for *da* in ENHG.

It is simpler to diagnose *thār* as an expletive than *da*, because it does not have the semantic ambiguity we have seen in ENHG (having not yet merged with the temporal adverb *thō*). Instead, if *thār* does not have a possible locative interpretation in a given clause, it can be assumed to be semantically null. In order to explore
the behavior of thār in these cases, I considered a small sample from the OHG Tatian text. This sample included 56 relative clauses, 41 of which had a subject gap. As predicted, many of the subject gap examples included an instance of thār, with no possible locative interpretation.

(21) bithiu uuanta mir teta mihhilu thie thār mahtīg ist because since to-me does much that THAR mighty is ‘Because the Mighty One did great things for me’ (Lucas 1:49)

(22) Thie thār habē ērun thie hōre. that THAR have ears that hear ‘He that has ears to hear, let him hear.’ (Matthes 13:43)

As in ENHG, thār occurs almost exclusively in the position of an extracted subject. Although the sample size is too small to say with complete confidence, this data allows a tentative estimate that thār occurred in subject gaps about 50% of the time in OHG.

I also compared the subject relatives in this sample to 35 parallel clauses from the Septembertestament which were also translated as a subject relative. The frequency of da/thār has significantly decreased between the two time periods (chi-square = 5.29, p = 0.021). This suggests that the use of the subject expletive has already begun to decline by the ENHG period.

According to this preliminary study of OHG, we can establish the following stages: in OHG, thār was available as an expletive, to optionally fill Spec,TP when
the subject is extracted. In ENHG, *da* shows the same property, although its status as an expletive is obscured by the semantic underspecification of the situation pronoun *da*. In Standard German, *da* is no longer available as an expletive in any environment. In the next section, I will propose an analysis for the historically attested occurrence of *da* in subject gap position, and show how it offers new insight into the status of the subject in the Germanic languages.

5 Analyzing *da* in Early New High German

From OHG until the present day, *da* patterns as a weak adverb, preferring a position in the left periphery of the verbal domain. In addition, while *da* is never used as a subject expletive in Standard German today, its OHG cognate thār ‘there’ did have this function. I have shown that ENHG can be taken as the middle stage in between these two systems: ENHG shows frequent use of the situation pronoun *da*, but also some evidence of a subject expletive *da* inherited from OHG.

Because of the ambiguous interpretation of the situation pronoun, I have shown that it is often difficult to distinguish semantically null cases of *da* from cases that are behaving as referential pro-forms for a less obvious situational context. However, the examples of *da* in subject gap position may occur in non-narrative and even ‘timeless’ contexts, in which there is literally no salient situation for it to refer to. When supported by the OHG evidence, this becomes clear evidence that subject gap *da* has the properties of a true expletive.

(23) Jhesu Christi, der *do* ist ein son Dauids des sons Abraham.
Jesus Christ who *DA* is a son of-David the son of Abraham
‘Jesus Christ, who is a son of David, Abraham’s son.’
(Septembertestament, Matthes 1:1)

(24) Selig sind, die *da* geystlich arm sind.
blessed are who *DA* spiritual poor are
‘Blessed are those who are spiritually poor.’
(Septembertestament, Matthes 5:3)

As Prince (1989) argued for Yiddish, I will propose that *da* occurs in Spec,TP when the subject is extracted from a position lower in the clause. However, unlike Yiddish, the appearance of *da* in subject gaps is optional (the evidence from OHG, in which the expletive occurs in roughly 50% of the subject gap clauses, supports such an assumption). My analysis of the ENHG expletive must hinge on a mechanism which allows subjects to optionally remain low.

Haeberli (1999, 2000, 2005) discusses two subject positions in the Germanic language family. Weak subject pronouns are restricted to the high position, as we have discussed above; full DP subjects, meanwhile, may optionally remain low. I follow Wallenberg (2009) in assuming these positions to be Spec,TP and the base
position of the subject in Spec,vP. Wallenberg also notes that the low subject position in German is related to the ‘definiteness effect’ discussed by Diesing (1992). Definite DPs, Diesing observes, prefer a high position; only indefinite subjects may (optionally) remain low.

I argue that in historical Germanic, expletive *da* is licensed when the subject is extracted from its base position in Spec,vP without first raising through Spec,TP. As Mikkelsen (2002) claims for *der* in Danish (Section 2.1), there is reason to believe that the gaps in subject relatives are inherently indefinite, regardless of the characteristics of their antecedents. Therefore, I argue that the subject of a relative clause behaves like an indefinite DP in German as well, and may optionally remain low (before extraction). In these cases, *da* is inserted to fill the empty position in Spec,TP.

This argument has the following consequences. First, the expletive *da* provides evidence that there is an EPP feature on T in German. When the extracted subject does not leave a trace in Spec,TP, the expletive is licensed to satisfy the EPP in its place. The evidence indicates that this feature changed at some point between OHG and contemporary German. This suggests that a null expletive became available to fill Spec,TP, and eventually replaced *da* in this function. As a result, *da* lost its expletive function, and behaves only as a situation pronoun in the modern language.

Second, the presence of *da* becomes support for the claim that there are two subject positions in German: the expletive may only occupy Spec,TP if the extracted subject can be licensed in a low position. Otherwise, we expect the subject trace to fill Spec,TP. We cannot easily motivate an analysis in which *da* is a pronounced version of the subject trace, because subject gap *da* is optional. By relating expletive *da* to the low subject position in German, we account for its optionality: indefinite subjects may be extracted without passing through Spec,TP. In this case, an expletive is licensed to fill the empty position. The subject may also move to Spec,TP, making the insertion of *da* unnecessary and impossible.

### 6 Conclusion

Understanding the behavior of *da* in ENHG has required not only that we compare the facts in contemporary German, but also that we look deeper into the past. By considering *da* in all attested stages of the language, we find that it has remained quite stable in some respects, but changed in others.

The exploration of this phenomenon has led to a deeper understanding of the subject position in German. The proposal of two subject positions in Germanic is supported by the behavior of *da*, which may only be licensed if the subject is extracted from its base position, and does not move through Spec,TP. In addition, we have seen evidence that the EPP in German behaves much as it does in languages like English.

This study relied on evidence drawn from the *Septembertestament* corpus of
Early New High German, as well as some pilot research using a small corpus of Old High German texts. This shows the importance of parsed corpora to empirical syntactic research. By using tools such as these, we come to understand synchronic facts of a language in ways that are not evident when viewed in isolation.

References


