On the historical development of pronouns referring to situations: the case of (expletive) *da, es* and *das*

Eric Fuß (Ruhr-Universität Bochum) & Roland Hinterhölzl (Università Ca' Foscari Venezia)
Sentence Grammar and Discourse Grammar IV, Oslo, 25.10.2019

1. Introduction

- **Background:** All present-day Germanic V2 languages exhibit Vorfeld-expletives when
  i. there is either no potential VF-filler available (as in impersonal passives);
  ii. other potential VF-fillers stay in a lower position for information-structural reasons:

  (1) a. *Es* ritten drei Reiter durch das Tor hinaus. (German)
      expl rode three riders through the gate out
  b. *Það* voru blöð, blek og pennar á borðinnu. (Icelandic)
      expl were paper ink and pens on table
  c. *Det* er løver i Afrika. (Norwegian)
      expl are lions in Africa
  d. *Er* is hier veel sneeuw. (Dutch)
      expl is here much snow

- VF-expletives are confined to the initial position of root clauses; they cannot occur in embedded clauses or in cases where the VF is already filled (overtly/by an OP):

  (2) a. *Es* wird getanzt. [impersonal passives/intransitive verbs]
      it is danced
  b. weil (*es) getanzt wird
      because it danced is
  c. Heute wird (*es) getanzt.
      Today is it danced

  (3) a. *Es* kamen drei Männer hinein. [thetic/presentational sentences]
      it came three men in
      ‘Three men came in.’
  b. Kamen (*es) drei Männer hinein?
  c. dass (*es) drei Männer hinein kamen.

The historical development of the VF-expletives – what we know

- Different elements in the various languages; still, they have some properties in common:
  i. **anaphoric** elements: third person singular neuter pronouns or locatives
  ii. short, monosyllabic forms
  iii. not attested in the earliest records/stages.

---

1 *er* is the weak form of *der*, the unstressed variant of the locative element *daar* ‘there’. But note that the locative *er* is homophonous with the partitive *er*, which goes back to a former genitival 3pl pronoun.
2 Similar developments can be observed in a couple of Romance languages (Old French *il*, in particular); however, in contrast to Germanic, the expletive elements are not confined to clause-initial position (subject expletives, but with special properties, e.g. use with impersonal passives: *il a été dansé*)
3 Note that the relevant forms in Scandinavian go back to demonstratives. However, it can be shown that *det/Það* first became part of the paradigm of personal pronouns as ‘new’ 3sg neut. forms before they turned assumed expletive functions (cf. e.g. Brugmann 1917).
• Earlier stages of German:
  (i) OHG: No VF-expletives (e.g., V1 order in existential/presentative clauses), cf. (4);

(4) uuarun thô hírta In thero lantskeffi uuahante [...] were then/there shepherds in that country abiding
Lt. *Et pastores erant In regione eadem. uigilantes [...] ‘And there were shepherds abiding in the fields [...]’
(Tatian, 85,29; Lk, 2,8)

(5) Ez ist geborn der heilant der die werlt alle erlosen sol it is born the Saviour who the world all deliver should ‘The Saviour was born, who should deliver the world...’
(12,1; Londoner Predigt(fragment), 39va,9-10)

**What we think we know**


• Problem: Chronological gap between the development of (i) a rigid V2 syntax and (ii) VF-expletives in the history of German.

• Potential solutions:
  i. no biconditional relation between V2 and Vorfeld-expletives
  ii. early German (from late OHG to mid/late MHG) is not a generalized V2 language
  iii. before the rise of Vorfeld-es, other elements were used as Vorfeld-expletives

• **Approach in terms of (iii)** (Fuß 2009, 2018): The rise of a rigid V2 syntax was accompanied by the development of the VF-expletive do/da (<<< deictic temporal/locative adverbs *tho/do* ‘then’ and *thar/da* ‘there’); see also Donhauser & Petrova (2009); cf. Light (2010, 2015) on da as a subject expletive in ENHG:⁴

(6) a. OHG: development of generalized EPP in C is accompanied by the rise of VF-da ⇒
   b. MHG: development of VF-es ⇒
   c. ENHG: VF-es is replaced by VF-da (cf. Fuß 2018)

---

⁴ Expletive *do/da* is still found in present-day dialects (cf. Mayerthaler & Mayerthaler 1990: 406, Weiß 1998):

(i) Do/*Es is gsunga und gantz worn.
   *EXPL is sung and danced been
   ‘There was singing and dancing.’
What we don’t know
- Open question(s):
  1. MHG: emergence/actuation of VF-es
  2. MHG/ENHG: Why was do/da replaced by the new VF-expletive es?

This talk:
(i) Closer look at the development of VF-es in MHG
(ii) New perspective on the pragmatic function(s) of (different types of) VF-‘expletives’
(iii) New hypothesis on the rise of VF-es that links its development to the changing role and text-linguistic function of VF-elements in German (text cohesion...; cf. e.g. Fleischman 1990)
2. The historical development of VF-es – previous work

- Widely accepted view on the development of expletive es:\(^5\)
  
  1. OHG: es is confined to the following functions:
     - personal pronoun
     - quasi-argument: weather verbs, impersonal constructions (experiencer verbs etc.)
     - correlate to extraposed clauses (and DPs...?)
  
  2. mid/late MHG: development of VF-es with impersonal passives (intransitive verbs) and thetic/presentational clauses.\(^6\)

Explanations for the rise of VF-es

- Brugmann (1917): Analogy with/reanalysis of correlate es (possibly under contact with Old French).
- Impersonal passives (intransitives): without es in OHG; development in MHG on the model of correlate-es in passives of transitive verbs with extraposed clausal complements:

  \(7\) eʒ ist ouch verboten von gehorsam, daʒ ...
  
  it is also prohibited by obedience that...
  
  (Berth. I, 530,22; Brugmann 1917: 30)

- Brugmann cites the following example from the Nibelungenlied (VF-es; no impersonal passive, though)

  \(8\) ez-n wart nie geste mēre baz gepflegn
  
  it-not were never guests more better treated
  
  ‘Never again were guests better treated.’
  
  (Nib. C, 698; Brugmann 1917: 30)

- According to Brugmann (p. 30), this development was promoted by the use of clause-initial da/do with impersonal passives:

  „Unterstützt wurde das Aufkommen dieser Ausdrucksweise dadurch, daß beiderseits, bei den zu Intransitiva und den zu Transitiva gehörigen Impersonali, von alter Zeit her im Eingang vielfach proklitische Partikeln, wie da, im Gebrauch gewesen sind: da wurde gekämpft wie da wurde gesagt, daß ...“

- Thetic clauses: According to Brugmann (1917: 34) the use of VF-es is already quite common in MHG, cf. (9) taken from Brugmann, and the examples in (10):

  \(9\) eʒ was ein küneginne gesezzen über só
  
  it was a queen sitting over sea
  
  (Nib. 49,5; Brugmann 1917: 34)

---


\(^6\) Haiman (1974) and Lenerz (1985) assume that this ‘topic-es’ may be reanalyzed as a subject expletive, which widens the distribution of expletive es and eventually leads to the loss of the distinction between VF-es and subject expletive es (as e.g. in Mainland Scandinavian).
(10) a. Iz giengen zwei mensch in ein goteshvos vnd baten da
it went two men in a god-house and prayed there
(Mettener Predigtsammlung)

b. Ich wil vich ein glichnisse sagen.
I want you a parable tell
Es giengen ze einen male ein fuchs vnd ein katze mit einander.
it went once a fox and a cat together
(Nikolaus von Straßburg: Predigten (C))

• Possible reasons for the development of VF-es in thetic clauses (Brugmann 1917: 35f.):
  ▪ Clausal typing (es as a means to signal the difference between questions and declaratives, see also Lenerz 1985: 124)
  ▪ Analogy with correlate es (extraposed clauses) and impersonal predicative constructions:

(11) Es schmerzt mich, dass ich wund bin ⇒ Es schmerzt mich meine Wunde
it hurts me that I sore am it hurts me my wound

(12) Es wurde Abend.
it became evening

  ▪ Reanalysis of correlate es in connection with extraposed (free) relatives that replace a subject DP:

(13) Es ist (alles) richtig, was du sagst.
it is all correct what you say

  ▪ generalized V2:

„Es kommt hinzu das Bestreben, das sich schon vor dem Aufkommen des syntaktischen es eingestellt hatte, die Anfangsstellung des Verbums mit einem proklitischen Wörtchen zu decken. Hierfür war es das gegebene Wort.“

  ▪ Discourse function: marking the introductory sentence of narratives (formerly marked by OHG thô, MHG dô or V1 order):


  ▪ More recent functionalist approaches (Vogel 2006, Czieza 2010): After the loss of V1-orders, VF-es evolved as a new means to mark thetic sentences, or, more generally, sentences with a rhematic subject

• Note: The above ideas are based on at best impressionistic evidence; the availability of large searchable corpora offers an opportunity to take a closer look at the relevant changes, with the possibility of unearthing new (quantitative) evidence and gaining a better understanding of the relevant historical developments.
3. Vorfeld-es in MHG: A corpus study

- Corpus study (MHG Reference Corpus (ReM); work in progress): Extraction of all cases where es/do/dar
  (i) directly precede the finite verb, and
  (ii) are directly preceded by a sentence boundary.\(^7\)
- Results (including doublets and annotation errors): 8,988 hits –
  - 5,582 instances of clause-initial do/da (including 121 cases of the pattern do-ne-Vfin)
  - 1,637 instances of dar (including 88 cases of the pattern dar-ne-Vfin)
  - 1,769 instances of es (including 254 cases of the pattern es-ne-Vfin)
- Focus of this talk: development of VF-es
- After elimination of doublets, annotation errors, and instances of subject/quasi-argument-es and subject/personal pronoun-es, 444 clear cases of VF-es remain.

3.1 Results

- 444 cases of VF-es (of 1,769 cases with clause-initial es; 25,1%)
- Distribution across time (absolute numbers):\(^8\)

![Figure 2: VF-es in ReM – distribution over time](image)

- The factor DEFINITENESS: 271/444 (61%) cases exhibit an indefinite subject
- Detail study: alles – it seems that the indefinite alles favors the use of a VF-es; in the whole MHG corpus there are only 2 examples where alles (3sg neut. nom.) appears in the

\(^7\) Compare the following query for VF-es and (ii) for cases with a negative particle intervening between es and the finite verb.

\[(i)\] \(\text{punc}=.E/ & \text{lemma}="\text{ër}" & \text{inflection}=\text{/Neut.Nom.Sg.3/} \& \text{pos}=\text{/VVFIN|VMFIN|VAFIN/} \& \#1 \cdot \#2 \& \#2_=_\#3 \& \#2 \cdot \#4\)

\[(ii)\] \(\text{punc}=.E/ \& \text{lemma}="\text{ër}" \& \text{inflection}=\text{/Neut.Nom.Sg.3/} \& \text{lemma}="\text{ne}" \& \text{pos}=\text{/VVFIN|VMFIN|VAFIN/} \& \#1 \cdot \#2 \& \#2_=_\#3 \& \#3 \cdot \#4 \& \#4 \cdot \#5\)

\(^8\) For the purpose of clarity, we slightly modified the annotated time information to fit the data into intervals of 50 years. In cases where ReM gives dates like “12,2-13,1”, we consistently chose the earlier value; in some cases, where ReM fails to include relevant information (as e.g. in the case of the Iwein), or deviates from its standard format “1x,1/2” (as in the Nürnberger Stadtbuch), we added relevant data.
Vorfeld (both from the first half of the 14\textsuperscript{th} century), but there are 38 cases where VF-es combines with a lower nominative indefinite pronoun \emph{alles}.

- In MHG, \emph{alles} is ambiguous between an indefinite pronoun (nom/acc 3sg neut.) ‘everything’, a quantifier ‘all’, and an adverb meaning ‘all, fully, completely, wholly’.
- In the examples that we counted as instances of VF-es, \emph{alles} is tagged as an indefinite pronoun in ReM; but still, many of them are ambiguous. Consider the following example from the \emph{Rolandslied} (second half of the 12\textsuperscript{th} century):

\begin{enumerate}
\item Iz was allez ueloren
  \begin{enumerate}
  \item ‘Everything is lost’
  \item ‘It was all/completely lost’
\end{enumerate}
\end{enumerate}

\begin{enumerate}
\item dass (es) alles neu ist
  \begin{enumerate}
  \item ‘that all is new’
  \item ‘that it is all new’
\end{enumerate}
\end{enumerate}

\begin{enumerate}
\item \textbf{Hunch}: Examples with \emph{alles} provided another context (in addition to correlative constructions) where \emph{es} (originally an argument) could be reanalyzed as a VF-expletive.
\item The factor \textbf{GENRE}: 170/445 (38,2\%) cases appear in legal texts (statute books and codes of law), which constitute only a minor portion of the overall corpus.
\item VF-es is particularly frequent in certain legal statements (requirements, prohibitions and commandments), compare the following examples:

\begin{enumerate}
\item a. Ez sol auch niemen chein silber verkauffen in der Stat.
  \begin{enumerate}
  \item ‘Nobody should sell silver in the city.’
\end{enumerate}
\item b. Ez sol auch ain icelich pecke alle tage nev pachen prot veile haben
  \begin{enumerate}
  \item ‘Each baker should have new bread for sale every day.’
\end{enumerate}
\item c. Ez sint verboten lose semeln [...] it are prohibited loose rolls
  \begin{enumerate}
  \item ‘Loose rolls are prohibited [...]’
\end{enumerate}
\end{enumerate}

\item This becomes particularly clear if we calculate the frequency of VF-es relative to text size (token number).\textsuperscript{9}

\textsuperscript{9} Note that we considered only texts which exhibited at least five instances of VF-es.
8

Figure 3: VF-es/10,000 words

- **Observations:**
  - The five texts with the largest share of VF-es are statute books/codes of law, mostly from the early 14th century.
  - VF-es is generally rare in narrative texts; a noteworthy exception is *Iwein*, which is a rather early text (around 1200) but still exhibits a share of VF-es which is only surpassed by later texts/texts from other genres.
  - Roughly one half of the early texts were adapted from Old French/Franco-Provençal (*Rolandslied*, 12; 13, 1: *Iwein, Straßburger Alexander*; of the 35 cases of VF-es in 12th-century texts, a large portion (14) is from Pfaffe Konrad’s *Rolandslied*).
- **VF-es vs. do:** first impressions based on a rough comparison of the above texts...\(^\text{10}\)

---

\(^\text{10}\) The numbers given for *do* in Figure 3 refer to all cases where *do* appears in clause-initial position (directly preceding the finite verb); in particular, no distinction is made between adverbial and expletive uses of *do* (telling apart the various functions of initial *do* is a difficult and time-consuming task, which we hope to advance in future work).
Figure 4: VF-es vs. do/10,000 words

- **dô vs. VF-es**: dô is typically used in narrative texts, but close to absent in legal texts (and treatises); there is only a single case in the *Nürnberger Stadtbuch*, but here, dô does not seem to be a text-continuative marker, but rather a locative (?) introducing an existential statement:

(17) wer daz brichet der gibet ie von dem pâr schuohe ij. schillingn.
who that breaks the-one gives each time of the pair shoes ij Shilling
dô sîn meister über gesetzt di daz besehen unde rüegen sulen.
DO are masters over set who that check and punish should
‘There are masters in charge who should check and punish that.’
(14,1; *Nürnberger Stadtbuch*, 8va,21-25)

- **Conclusions**:  
  - The traditional view that VF-es emerged in 13\textsuperscript{th}/14\textsuperscript{th} century MHG is basically correct.  
  - However, there are also a couple of earlier cases dating to the 12\textsuperscript{th} century.  
  - There are some indications that contact with Old French might have played a role (see also Brugmann 1917), but as yet the evidence is not conclusive;  
  - VF-es is particularly frequent in legal texts; in narrative texts it is rare;  
  - dô is frequently used in narrative texts, but rare in legal texts (and treatises);  
  - Discourse functions – es vs. dô:  
    - dô: discourse-continuative marker that links its clause to a tense interval that has been established in the previous discourse (see also Trips & Fuß 2009 on OE);
VF-es: used in clauses that (i) are temporally independent (not bound to the previous discourse context), or (ii) introduce a new tense setting/situation.

- The distribution of VF-es across MHG texts (in particular, the contrast between narrative and legal texts) suggests the following trajectory for the spread of VF-es:
  - in MHG, VF-es became popular as a means to introduce certain legal statements (non-narrative texts), which were typically produced by the mediaeval chanceries;
  - when the language use of the chanceries became a model for written German in the ENHG period, the VF-es spread to other (i.e., narrative) text types.

4. Towards an analysis: On the discourse function of (VF-) expletives

4.1 Some initial observations

Temporal Anchoring: the role of a reference situation

(18) a. John visited his mother
    b. \(\exists e \text{ visiting(e) \& past (e) \& agent (e, John) \& theme (e, his mother)}\)

(19) a. John visited his mother. (e_1)
    b. She was sick (e_2)
    c. \(e_1 < e_2 < s, e_2 < e_1 < s, e_1 \circ e_2 < s\)
    d. She was sick one week before/later

Tense is treated as a predicate on times or intervals, that is to say, it is assumed to have temporal arguments (cf. Stowell 1993 and much subsequent work). Instead we must assume that Tense expresses relations between situations / events.

(20) Past \((s_1, s_2) = s_1 \text{ precedes } s_2 =: \tau (s_1) < \tau (s_2)\)

Weak and strong definites

(21) a. A man came in. The man wore a green hat.
    b. A person was killed in a terrible way. The murderer must be insane

(22) a. Hans hat gestern Maria getroffen. Er hat seine Freundin sofort umARMT
    b. Hans hat gestern Maria getroffen. Er hat sofort seine FREUNDin umarmt.
    John has yesterday Mary met. He has immediately embraced his girlfriend

Conclusion: the use of temporal predicate in (13) is anaphoric

Germans shift the deictic center rather than make an adequate use of proximal and distal deictics:

(23) Ich machte letztes Jahr Urlaub in Island. Da regnete es (in Iceland, not at U)

We note that \(da\) has lost its original deictic use in which it locates an event with respect to the utterance situation. In Italian, (23) would sound weird, requiring the use of the distant deictic \textit{qua}.

(24) a. ??Ho fatto una vacanza in Islanda l'anno scorso. Qui pioveva (here rained)
    b. Ho fatto una vacanza in Islanda l'anno scorso. Qua pioveva. (there rained)
The pronoun *es* always seems to link the referred to event of the main predicate back to the utterance situation, rather than linking it to a pre-established reference situation in the context, as is illustrated in (25).

(25) ??Ich machte letztes Jahr Urlaub in Island. Es regnete. (some time in the past)

This implies that *es* has probably replaced *da* in its deictic use. In particular *es* introduces a new reference situation, while *da* picks up one established in the context. This means that *es* has replaced *da* in its post-finite use, as in (26).

(26) warun tho hirta in thera lantskeffi

But before *es* enters the picture V2 is generalized producing structures like (27), in which *tho* functions as an explicative; it makes explicit a discourse connection that was expressed by V1-order alone before

(27) tho warun hirta in thera lantskeffi

*Tho* in (27) indicates that tense has an anaphoric interpretation (it picks up an already introduced situation in the past).

### 4.2 *es* as quasi-argument and as Korrelat

What can be said about the status of *es* as a quasi-argument with weather verbs? If we agree that *(es) regnete* denotes the set of situations (in the past) in which it rains, then weather verbs have to be assumed to combine with a situation argument to arrive at a truth value:

(28) $Es$ regnete / It rained
    $s_t$ (that is identified with the reference situation) $\in \{ s \mid \text{rains in } s \}$

In (28), *es* expresses that a new situation that is identified by tense and linked to the utterance situation is an element of the set of situations in which it rained - a clear case of predication.

The only difference to predications like *John sang* is that in the former case the argument is of the situation type, while in the latter case the argument is of the individual type. The second difference concerns the fact that the argument of weather verbs is introduced by tense, that is to say, by a temporal relation, while individual arguments are introduced via theta relations.

Again, Tense must be taken to express a relation between two situations, one identified with the utterance situation and the other functioning as the reference situation, the latter of which is arguably lexicalized by *es*.

Is there independent evidence for this assumption? Other predicates that only have a subject argument that is not of the individual type but a full clause, like *eine Schande sein, oder mich freuen* display *es*, if the clause is extraposed (cf. 29cd). In this case, *es* is treated as a Korrelat.

(29) a. Dass Peter zu Hause blieb ist eine Schande
    That Peter remained at home is a shame

b. Dass Peter zu Hause blieb freute mich
    That Peter remained at home pleases me

c. Es ist eine Schande dass Peter zu Hause blieb
    It is a shame that Peter remained at home
d. Es freute mich dass Peter zu Hause blieb.
   It pleases me that Peter remained at home

We can equate the use of (29ab) with the use of (29cd) if we assume that the clauses in (29ab) denote a given situation. Note that a proposition is normally taken to denote a set of situations. I assume here that a proposition anchored to the context denotes a single situation. In (29ab) case, the situation that Peter remained at home (and known to the hearer) is asserted by the speaker to be a shameful situation / a situation that arouses joy.

*es* in (29cd) introduces a new situation that is characterized as one in which it is judged as shameful or joyful the known or new proposition that Peter remained at home. In other words, in (29ab) the assertion is anchored via a situation topic to the context. In (29cd), the utterance has a default anchor.

This interpretation of *es* in (29) is corroborated by the observation in (30): *es* cannot function as a discourse anaphor. *Es* is the weak variant of the strong pronoun *das* referring anaphorically to a given situation.

(30) a. Peter blieb zu Hause. Das / ??es ist eine Schande
   b. Peter blieb zu Hause. Das / ??es freute mich

Thus, we have a pair of expressions *das* and *es* that represent the strong and weak determiner of a nominal denoting a situation: the strong one is discourse anaphoric, the weak one establishes a new discourse referent. In other words, *es* has an existential impact on the assertion of the speaker, as is indicated in (31).

(31) a. Es regnete = there is a situation in the past in which it rained
   b. Es gab einen Aufruhr = there is a situation s in the past & there is an x such that
      Aufruhr (x) in s

The concrete proposal is that *es* is a generalized quantifier combining with an abstract noun denoting a situation and a predicate on situations (a property of situations). *es* can be taken as the spell-out of a D-head that selects a nominal denoting a situation. As a weak determiner, it can be identified with the t-operator deriving the reading specified in (32b) for the sentence in (32a).

(32) a. Es regnete
   b. t s.  s < U & rain in s (there is a unique s preceding the utterance time in which it rains)

The analysis of *es* as a generalized quantifier predicts that *es* combines with a tensed verb in T in a strictly compositional fashion: first the meaning of T and vP is computed, then the generalized quantifier in [Spec,TP] combines with the meaning of nP and T', as is indicated in (33).

(33) a. λs. s < U & rain in s (meaning of Tense +vP)
   b. t s.  λs λP (meaning of *es*)
   c. t s.  s < U & rain in s (meaning of TP)
What is the difference between weather and existential es versus Korrelat-es?

While the predicate is complete with Korrelat-es and the predication only needs to be anchored to the context in the absence of a given situation or individual, in cases of weather and existential es the predicate is incomplete and - because of the nature of the predicate - the resulting predication can only be anchored with the default anchor es. (29ab) thus resemble (34a) and (29cd) resemble (34b).

(34) a. Hubert Haider spricht (one of the speakers is Hubert Haider)
    b. Es spricht Hubert Haider (it speaks Hubert Haider)

COMMENT: Can this possibly used to provide a formal explanation for Brugmann’s assertion that VF-es evolved in analogy to/as a reanalysis of correlate-es?

5. Concluding summary and directions for future research

References

Axel, Katrin. 2007. *Studies on Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-Second*. Amsterdam: John Benjamins


