Reanalyzing German correlative *es*

Anke Holler

University of Göttingen

{anke.holler@phil.uni-goettingen.de}

1 Introduction

Several Germanic languages use correlatives to mark subordination. German is considered to be a prime example of a language realizing correlative constructions to embed finite argument clauses. This paper discusses complex clausal constructions in German that can be identified by a so-called correlative nominal element *es* ('it') and a right-peripheral full clausal argument that is linked to *es*. On an intuitive level, *es* functions as an antecedent of the linked argument clause in these constructions. Although correlative *es* is a well-established phenomenon of German grammar, there is no theoretical account that captures the empirical facts comprehensively. In particular, the janus-faced nature of *es* between an expletive and a referential realization form often remains unnoticed. In the present paper, this homonymy of correlate *es* is empirically substantiated. On the basis of these empirical observations the paper develops a new constraint-based analysis. The paper is organized as follows: After describing the phenomenon in section 2 and summing up the results of previous studies dealing with correlative *es* in section 3, empirical data that has not yet been captured in existing proposals are given in section 4. Taking into account the new data basis, section 5 then develops and outlines the aforementioned new constraint-based analysis. To conclude, the results of the paper are presented in section 6.

2 Phenomenon

The correlate *es* construction is characterized by a correlative *es* that occurs in the matrix clause and relates in some way to a finite *dass* ('that')-marked clause serialized to its right in the syntactic surface structure.[1] Most of the standard approaches assume that the *dass*-clause is located in an extraposed position since it follows the matrix clause's finite verb if this is linearized sentence-finally.

(1)   Hotzenplotz bedauert es, dass er außer Räuberei nichts gelernt hat.
       Hotzenplotz regrets it that he has nothing learned except robbery.

In this configuration *es* is usually analyzed as a means of recursive sentence embedding, which functions as a structural element filling a syntactic position and referring cataphorically to the right-peripheral argument clause.

As has been already observed in traditional grammar of German the occurrence of *es* is also subject to certain topological restrictions. Since German is a verb-second language, it offers a so-called prefield position. In main clauses, this results from fronting the finite verb. If the *dass*-clause is topicalized to this position, *es* is obligatorily omitted, cf. (2). Also, the *dass*-clause may not be serialized adjacent to *es* in the so-called middle field, cf. (3).

(2)   Dass er außer Räuberei nichts gelernt hat, bedauert (*es) Hotzenplotz.
       That he has nothing learned except robbery regrets Hotzenplotz.
(3)   […], weil Hotzenplotz (*es), dass er außer Räuberei nichts gelernt hat, bedauert
       […] *because Hotzenplotz regrets that he has nothing learned except robbery*

Any analysis treating the correlate *es* construction has also to cover these topological facts.

---

[1] In fact, there are further infinite construction types involving *es* that are not considered in this paper, cf. Müller (1999). The presented analysis, however, can easily be transferred to these types.

3 Previous Approaches

3.1 Standard generative approaches

In principle, there are two competing approaches to correlative *es* in research literature. They differ fundamentally in two respects: firstly, in the analysis of the syntactic and semantic status of correlative *es,* and, secondly, in the interpretation of the grammatical relation between *es* and the linked right-peripheral finite clause

One strand of research (e.g. Bennis 1987, Cardinaletti 1990, Sonnenberg 1992) analyzes *es* as a case- and theta-marked argument of the matrix verb. Correlative *es* projects a nominal phrase and contributes a referential index. The corresponding finite *dass*-clause functions as an explicative (appositive) attribute of *es*. This view is usually implemented by adjoining the *dass*-clause to a verbal projection (V' or VP) containing *es* as a verbal argument. One consequence of this analysis is that the correlative *es* and the finite *dass*-clause constitute two independent constituents.

The other strand of research holds that correlative *es* and the finite *dass*-clause form just one (discontinuous) nominal constituent (e.g. Zimmermann 1993, Zifonun 1995, Müller 1996, Sudhoff 2003, Sternefeld 2006) that is subcategorized and theta-marked by the matrix verb. In this constellation *es* behaves as an expletive, which is linked to the extraposed *dass*-clause. The specific approaches of this analysis variant differentiate w.r.t. the integration of the *dass*-clause into the nominal phrase containing *es*. Müller (1996) and Sudhoff (2003) propose that *es* acts as the functional head of this nominal phrase and obligatorily selects the *dass*-clause as its complement, while Zimmermann (1993) argues that the *dass*-clause should modify the maximal nominal projection.

3.2 Constraint-based approaches

Although the role of correlative elements in recursive sentence embedding has received considerable attention in German grammar writing and the last decades' generative theory, only a few constraint-based approaches have dealt with the issue so far. The two most prominent constraint-based approaches are those by Berman et al. (1998) and Kathol (1995). Starting with Berman's approach, the two approaches can be outlined as follows:

Berman et al. (1998) develop an unification-based analysis in the framework of lexical-functional grammar that relies on two basic assumptions: (i) *es* behaves like a referential pronoun, and (ii) *es* and the *dass*-clause share the same argument slot of the matrix predicate, which is achieved by unifying their f-structure contributions under the same function. Consequently, both *es* and the *dass*-clause differ at the categorical level (*es* is analyzed as an NP, the *dass*-clause as a CP) but share the same grammatical function OBJ at the level of grammatical functions. The unification analysis is charming since the *dass*-clause does not have to be categorized syntactically as an appositive or adjoined clause although *es* is interpreted referentially. Moreover, the co-occurrence of correlative *es* with a *dass*-clause is licensed by general constraints on c-structures and f-structures in a LFG fragment of German. Semantically, however, the proposition introduced by the *dass*-clause restricts the independently introduced variable of the referential pronoun *es* by adding more information. Hence, the *dass*-clause behaves semantically like a typical apposition. In this respect, their approach follows the assumptions of standard generative approaches analyzing *es* as a referential pronoun. The main criticism of Berman's approach is that it overlooks empirical data showing that *es* if it occurs with certain verbs may also behave like an expletive. The set of data substantiating this criticism is given in section 4 of this paper.

Kathol's (1995) HPSG-based proposal for the analysis of correlative *es* shares with the presented LFG-analysis the assumption that *es* has properties of a referential pronoun. Hence, Kathol explicitly criticizes Pollard and Sag's (1994) treatment of similar constructions in English. Pollard and Sag analyze *it* as an expletive form that does not bear a semantic role. In contrast to this, Kathol reverses the relationship between syntactic complements and their semantic representations in correlative

constructions by arguing that "it is the pronominal that bears the thematic role previously thought to be borne directly by the propositional element. The latter in turn is not a direct semantic argument of the predicate in question, but instead is linked to the role assigned to the index of es" [p. 289f]. Kathol establishes this linkage via a CONTEXT feature ANCHOR, which takes two arguments: the restricted nominal index of *es* and the index of the correlated clause being of type *parameterized states-of-affairs*. This mechanism is necessary to avoid a type clash. A type clash would be the consequence if both indices were structure-shared directly. In addition, Kathol formulates a constraint saying that a constituent whose content value is linked via an ANCHOR relation to the index of some other entity is required to occur extraposed in order to account for the above-mentioned topological facts.

In comparison to the approach of Berman et al., Kathol does not act on the syntactic level of grammatical functions but on the semantic level of argument structure. This is, however, a negligible difference. More important is the fact that both approaches suffer from the lopsided view on *es* as a referential pronoun. That *es* cannot exclusively be interpreted as a referential pronoun is suggested by data like (4) taken from Frey (2011).

(4)   a. Was lehnt Maria (*es) ab?   *What did Maria refuse (*it)?*
       b. Was hat (*es) Maria lange geglaubt?   *What did Maria believe (*it) for a long time?*

The examples show that the *dass*-marked clause cannot be questioned if correlative *es* is realized. If *es* was a referential pronoun in cases like (4), one would expect that it was possible to answer the questions with *es*.

3.3 Summing up

Neither of the existing analyses is able to account for German correlative *es* in its entirety. The main reason for this is that all previous approaches lack the generalization that correlative *es* may behave both as an expletive and as a referential pronoun depending on the respective syntactic context. In the following section we will present corpus-based support for the hypothesis that the German correlative *es* distinguishes between two types: an anaphoric referential pronoun and a true expletive. Pütz (1975) has already stated that these two types of correlative *es* may exist. His claim, however, is based on introspection and does not rest on empirical data.

4   Janus-faced *es*

In this section, we will argue on the basis of empirical data that correlative *es* is homonymous between an expletive and a referential form. A first step in proving this hypothesis is the evaluation of so-called correlate-taking verbs.

It is a well-established assumption of standard German grammar that verbs may be classified w.r.t. their ability to select correlative *es*. Surprisingly, there is no consensus in research literature about this classification. For instance, so-called verba dicendi and sentiendi like *sagen* ('say'), *meinen* ('think'), *hören* ('hear'), *behaupten* ('assert') etc. are sometimes ranked as correlate-taking and sometimes as correlate-rejecting. The list of inconsistently classified verbs could be extended. One reason for the uncertainty in the evaluation of the respective verbs may be that the empirical basis of the classifications is often very thin. The classifications often rely on construed examples or on unsystematically collected corpora. In the latter case a single item taken from a corpus is often regarded as sufficient evidence for a certain hypothesis. Boszák (2009) is a recent example of this fallacy.

Based on a quantitative corpus study[2] we can empirically substantiate that in fact there is a categorial distinction between two verb classes: With the first class of verbs, correlative *es* is robustly attested. This is shown in figure 1 where the blue bars indicate the number of examples with *es*.[3] With the second class of verbs, however, correlative *es* is not attested among the hundred examples investigated as can be seen in figure 2.
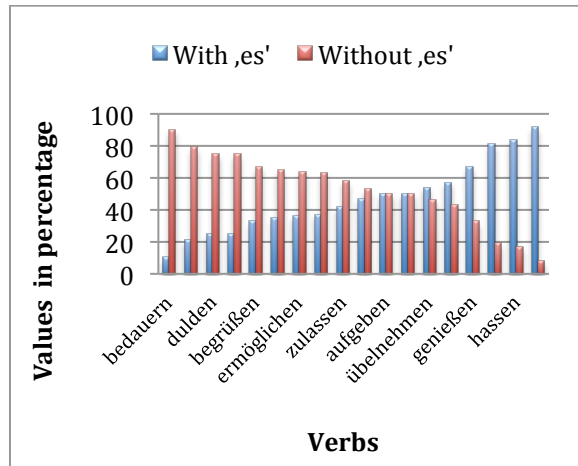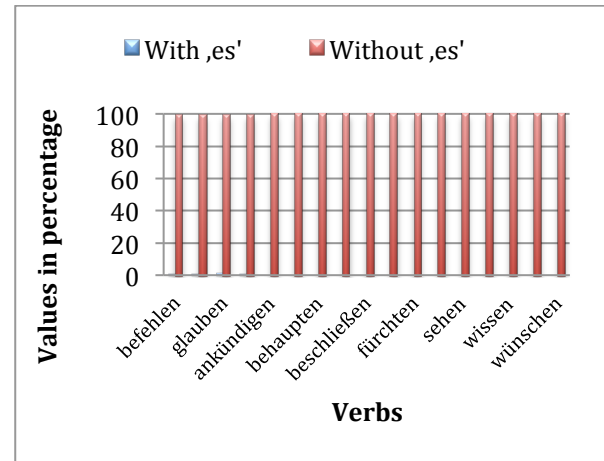

Figure 1: Verbs of class I

Figure 2: Verbs of class II

At first glance, the result for the second verbal class is incomprehensible since it comprises verbs like *glauben* ('to believe'), *sagen* ('to say') and *wissen* ('to know') that are usually regarded as correlate-taking in German grammar theory. In fact, sporadic examples like (5) where one of these verbs is used with correlative *es* can also be found in corpora, although the quantitatively obtained results for the second verbal class seem to be clear-cut.

(5)   Es ist schrecklich, wenn vor so vielen Dingen ein dunkler Vorhang ist. Ich möchte ihn immer nur zerreißen, aber ich kann es nicht. Ich glaube es dir, dass Du den Vorhang nicht zerreißen kannst. *It is awful that so many things are behind a dark curtain. I believe it that you cannot tear the curtain* [cited from GDS: 1487]

The puzzling empirical situation suggests examining the direct context of the *es*-containing complex clauses with a predicate of class II in more detail. As a result of this, one recognizes that in all of these cases *es* seems to refer back to a contextually given, discourse-old entity and hence behaves like an anaphoric element. Thus, the underlying reason for the divergent classification of verbs w.r.t. their correlate-taking ability in the literature is due to the Janus-faced nature of correlative *es*. Obviously, *es* occurs in two realization forms: Combined with verbs of class I it just fills a syntactic position and functions as a placeholder, which is a structural element without any semantic contribution; combined with verbs of class II, however, it must be analyzed as an anaphoric pro-form referring back to a pre-mentioned state-of-affairs. The presented corpus evidence supports introspective data by Pütz (1975), Sudhoff (2003) and Frey (2011), who conjecture on theoretical grounds that at least two classes of putative correlative-*es*-taking verbs need to be distinguished. These classes differ for example w.r.t. *wh*-extraction and V2-embedding facts. Moreover, Axel/Holler/Trompelt (t.a.) have shown in a psycholinguistic study that *es* may function as both a non-referential structural element and a

---

[2] Methodically, we selected a group of 35 verbs for which divergent judgments exist in the literature. For each verb, the number of hits was limited to 1000 by random selection. Of those 1000, we manually selected the first 100 examples in which the *dass*-clause is really the object clause of the critical verb.

[3] As can be seen, the verbs given in figure 1 are attested with *es* to varying degrees. The frequencies range from 10% with *bedauern* ('regret') to more than 90 per cent with *überlassen* ('leave to').

referential anaphoric pro-form. The study demonstrates that the respective usage depends on the syntactic contexts and the verbal class involved.

5 A comprehensive analysis for correlative *es*

Since existing theoretical approaches of correlative *es* overlook the above-mentioned homonymy of *es*, there is still a need for a comprehensive analysis accounting for the presented empirical facts. The respective analysis must be able to differentiate between a placeholder and a pro-form usage of *es*, and it must mark verbs w.r.t. their ability to occur with a placeholder. The latter can be reached by sub-sorting the HEAD value such that verbs can be divided into *placeholder-accepting* and *placeholder-non-accepting* verbs in the lexicon. In addition, *es* is lexically homonymous between an expletive and a referential pronoun.

Only in the first case, *es* belongs to the class of functional elements and hence its CONTENT value is instantiated by the empty set. Syntactically, however, functional *es* is subcategorized for a CP marked as + EXTRA, cf. Keller (1995), Kathol (1995). Consequently, together with its complement, i.e. the *dass*-clause, *es* projects a DP that is case- and theta-marked by the matrix predicate if this is of sort *placeholder-accepting*. Because of its positively specified EXTRA value it follows from general constraints on extraposition that the CP has to occur extraposed. Thus, the topological facts presented in section 2 are captured since a realized functional *es* forces the *dass*-CP to be positioned right-peripheral, which particularly means that it neither can be topicalized nor realized in the so-called middle field.

As a pro-form, *es* is analyzed as a lexical pronominal element, which is fully saturated and whose SUBCAT list is empty. Semantically, however it introduces an index and an anaphoric relation that relates referential *es* to its antecedent. Since referential *es* may introduce a saturated nominal phrase (i.e. a DP) on its own (ergo: independent of the co-occuring *dass*-clause), referential *es* is able to function as an argument of the respective matrix predicate, which means that its index instantiates one of the ARG values of the verbal relation introduced by this predicate. This value must be structure-shared with the first argument of the anaphoric relation. Thus, like Kathol's anchor relation, the anaphoric relation combines the referential anaphoric pronoun *es* with an entity of sort *parameterized state-of-affairs*. The right-peripheral *dass*-clause is analyzed as a non-integrated clause being of sort *fully-non-integrated* as argued in Holler (2008). From this the topological facts can be derived easily.

The syntactic status of the *dass*-marked clause hence depends on the realization form of *es*. If *es* functions as a placeholder, the *dass*-clause is a complement of *es*; if *es* functions as an anaphoric pro-form, the correlated *dass*-clause represents a non-integrated clause that behaves like an appositive (explicative) attribute. From this follows immediately its final position at the right edge.

6 Conclusion

In the present paper it has been argued that correlative *es* functions either as a placeholder which is a structural element without any semantic contribution, or as an anaphoric pro-form, which must be resolved by a suitable state-of-affairs to its left. It has been shown empirically that the placeholder vs. anaphoric use of correlative *es* is both verb-class dependent and context dependent. We developed a constraint-based analysis that accounts for the empirical facts and thus differentiates between *es* as an expletive and *es* as a referential pronoun. In the first case, *es* is analyzed as a functional element and in latter case *es* is treated as canonical pronoun. Correspondingly, the dependent *dass*-class behaves as a complement of *es* or as a non-integrated clause.

It is to be expected that the analysis presented here can be transferred to other languages possessing correlative elements such as Dutch and Italian. This should be examined carefully in further research.

References

Axel,K./Holler,A./Trompelt, H. (t.a.). Correlative *Es* vs. *Das* in German: An Empirical Perspective. In: Schwabe, K./Meinunger, A./Frey, W. (eds.), Inner-sentential propositional pro-forms, Amsterdam: John Benjamins.

Berman, J./Dipper, S./Fortmann, Ch./Kuhn, J. (1998). Argument clauses and correlative *es* in German – deriving discourse properties in a unification analysis. In Butt, M./King, S. (eds.) Proceedings of LFG98 conference, Stanford: CSLI Publications.

Bennis, H. (1987). Gaps and Dummies. Dordrecht/Providence: Foris Publications.

Boszák, G. (2009). Realisierung der valenzbestimmten Korrelate des Deutschen. Frankfurt/M. u.a.: Lang.

Cardinaletti, A. (1990). Impersonal Constructions and Sentential Arguments in German. Padova: Unipress.

Frey, W. (2011). On the ‚nominal' character of clauses associated with a pronominal. Ms., ZAS Berlin.

Holler, A. (2008). German dependent clauses from a constraint-based perspective. In Fabricius-Hansen, C./Ramm, W. (eds.) "Subordination" vs. "Coordination" in Sentence and Text. Amsterdam: John Benjamins,187-216.

Kathol, A. (1995). Linearization-based German syntax. PhD thesis. Ohio State University.

Keller, F. (1995). Towards an account of extraposition in HPSG. Proceedings of the seventh conference on European chapter of the Association for Computational Linguistics, 301-306.

Müller, G. (1995). On Extraction and Successive Cyclicity. In Lutz, U./Pafel, J. (eds.) On Extraction and Extraposition in German. Amsterdam/Philadelphia: Benjamins, 213-244.

Müller, S. (1999). Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche. Linguistische Arbeiten 394. Tübingen: Niemeyer.

Pütz, H. (1975). Über die Syntax der Pronominalform es im modernen Deutsch. Tübingen: Narr.

Sternefeld, W. (2006). Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen. Tübingen: Stauffenburg.

Sonnenberg, B. (1992). Korrelate im Deutschen. Tübingen: Niemeyer.

Sudhoff, S. (2003). Argumentsätze und es-Korrelate – zur syntaktischen Struktur von Nebensatzeinbettungen im Deutschen. Berlin: Wissenschaftlicher Verlag.

Zifonun, G. (1995). Minimalia grammaticalia: das nicht-phorische es als Prüfstein grammatischer Theoriebildung. Deutsche Sprache 23, 39-60.

Zifonun, G./Hoffmann, L./Strecker, B. (1997). Grammatik der deutschen Sprache. Berlin/New York: de Gruyter. [= GDS]

Zimmermann, I. (1993). Zur Syntax und Semantik der Satzeinbettung. In: Rosengren, Inger (Hrsg.)(1993). Satz und Illokution. Band 2.