Exhaustiveness effects in clefts are not truth-functional

Heiner Drenhaus a, *, Malte Zimmermann b, Shravan Vasishth b

a Department of Computational Linguistics & Phonetics, Saarland University, 66123 Saarbrücken, Germany
b Department of Linguistics, Universität Potsdam, Haus 24/35, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Article history:
Received 25 May 2010
Received in revised form 26 October 2010
Accepted 28 October 2010

Keywords:
ERP
It-clefts
Only-foci
Information structure
German

While it is widely acknowledged in the formal semantic literature that both the truth-functional focus particle only and it-clefts convey exhaustiveness, the nature and source of exhaustiveness effects with it-clefts remain contested. We describe a questionnaire study (n = 80) and an event-related brain potentials (ERP) study (n = 16) that investigated the violation of exhaustiveness in German only-foci versus it-clefts. The offline study showed that a violation of exhaustivity with only is less acceptable than the violation with it-clefts, suggesting a difference in the nature of exhaustivity interpretation in the two environments. The ERP results confirm that this difference can be seen in online processing as well: a violation of exhaustiveness in only-foci elicited a centro-posterior positivity (600–800ms), whereas a violation in it-clefts induced a globally distributed N400 pattern (400–600ms). The positivity can be interpreted as a reanalysis process and more generally as a process of context updating. The N400 effect in it-clefts is interpreted as indexing a cancelation process that is functionally distinct from the only case. The ERP study is, to our knowledge, the first evidence from an online experimental paradigm which shows that the violation of exhaustiveness involves different underlying processes in the two structural environments.

© 2010 Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: +49 681 302 6597; fax: +49 681 302 6560.
E-mail addresses: h.drenhaus@gmail.com (H. Drenhaus), mazimmer@uni-potsdam.de (M. Zimmermann), vasishth@uni-potsdam.de (S. Vasishth).

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
1. Introduction

There is a general consensus in the semantic literature that not only sentences with the exclusive focus particle only (1 a), but also it-cleft sentences (1 b) come with an exhaustive interpretation; see, for instance, Halvorsen (1978; 1976), Atlas and Levinson (1981), Horn (1981), Declerck (1988), Delin and Oberlander (1995), É. Kiss (1999; 1998), Hedberg (2000), Krifka (2008):

(1) (a) Only [John] FOC stole a cookie.
    (b) It’s [John] FOC that stole a cookie.

As in the only-focus sentence in (1 a), the cleft sentence in (1 b) seems to say that nobody else but John stole a cookie. Due to their exhaustive interpretation, both only-foci and focus it-clefts have been argued to be incompatible with certain additive particles, such as postposed too, cf. (2):

(2) (a) *Only [John] FOC, too, stole a cookie.
    (b) *It’s [John] FOC, too, that stole a cookie.

While the existence of an exhaustiveness effect in (1a) and (1b) is beyond doubt, it is not so clear that the effect arises for the same semantic or pragmatic reasons. Specifically, while it is generally acknowledged that the exhaustiveness effect in (1a) forms part of the asserted truth-functional content of the utterance (Beaver & Clark, 2008; Horn, 1969; König, 1991; Rooth, 1985), various semantic and pragmatic explanations have been proposed in order to account for exhaustiveness in it-clefts. While some of the proposed solutions have been rejected on theoretical grounds, others remain viable options, with the result that the issue has not been conclusively settled among theoretical linguists.

The present contribution aims at shedding light on the ongoing theoretical debate from an experimental perspective. We provide experimental evidence that the violation or cancelation of exhaustiveness effects induce distinct processing costs in different structural contexts. We do so by presenting the results of a comparison of ERP patterns in the processing of (marked) violations of exhaustiveness in semantically exhaustive only-foci, on the one hand, and in the processing of (marked) violations of exhaustiveness with German it-clefts, on the other. Our findings pertain to the larger question of how the human language processor responds to the different types of demands initiated by the tested structures.

2. Theoretical background

This section briefly lays out the theoretical background assumptions for our study. Section 2.1 gives a brief overview of semantic analyses of exhaustiveness effects with only-foci and it-clefts, pointing out which analyses remain viable options. Section 2.2 summarizes some of the major eyetracking studies on focus and it-clefts, and Section 2.3 introduces some relevant background information on ERP-studies. Following this, we present the experimental design and hypothesize on some possible outcomes in section 2.4.

2.1. The analysis of exhaustiveness in only-foci and it-clefts

There is a general consensus that exhaustiveness with only-foci enters the semantic representation as part of the asserted truth-functional content. This can be seen, for instance, by the fact that exhaustiveness can be targeted by negation, cf. (3a). Given this, the truth-conditions of (1a) in terms of universal quantification are given (in simplified form) in (3b):

(3) (a) [[ Not only Bill stole a cookie ]] = 1 iff Bill is not the only one that stole a cookie. = somebody besides Bill stole a cookie.
    (b) stole_a_cookie(john) ∧ ∀z [ stole a cookie(z) → z = john]
For *it*-clefts, by contrast, the situation is not as simple. Based on empirical results and theoretical considerations, the semantic literature offers four positions on the semantic or pragmatic nature of the exhaustiveness effect in (English) *it*-clefts. It has been alternatively analyzed as a

(4) (i) truth-functional effect (on par with only-foci) (Atlas & Levinson, 1981; É. Kiss, 1999; 1998);
(ii) conventional exhaustiveness implicature of the form nobody but, which is added to the asserted content (Halvorsen, 1978; 1976);
(iii) uniqueness or maximality presupposition ((Delin & Oberlander, 1995; Percus, 1997; Wedgwood, Pethö, & Cann, 2006; von Fintel & Matthewson, 2008); or a
(iv) (generalized) conversational implicature (Horn, 1981).

The different analyses make different empirical predictions concerning the behavior of *it*-clefts. For instance, we would expect *it*-clefts and their corresponding only-focus constructions to be synonymous on the truth-functional account (4i), cf. (1a and 1b). Conversely, we would expect *it*-clefts and their corresponding only-focus constructions to behave very differently on the conversational implicature account in (4iv). For instance, we would predict the exhaustiveness effect to be cancelable with the *it*-cleft in (1b), but not with the only-focus in (1a).

Horn (1981) provides strong arguments against an analysis of the exhaustiveness in *it*-clefts in terms of (4i) (truth-functional) or (4ii) (conventional implicature). First, Horn demonstrates that *it*-clefts and only-foci are in fact not synonymous, effectively ruling out (4i). Second, he shows that the exhaustiveness effect in *it*-clefts does not show the expected projection behavior of conventional implicatures (see, e.g. Karttunen & Peters, 1979) in negative, question, and conditional contexts, thus ruling out (4ii).

The remaining candidates (4iii) (presupposition) and (4iv) (conversational implicature) fare equally well with respect to a standard test for non-asserted (i.e. non truth-functional) content, the ‘Hey, wait a minute!’-test (von Fintel, 2001). In (5a), the exhaustive meaning component in A’s *it*-cleft can be questioned or suspended by B’s ‘Hey, wait a minute!’-response, whereas the same is not possible for the exhaustive component in only-foci (5b).

(5) (a) A: It’s Peter that took part in the contest.
   B: Hey, wait a minute! I thought there were many participants.
(b) A: Only Peter took part in the contest.
   B: #Hey, wait a minute! I thought there were many participants.

However, the presupposition account and the conversational implicature account differ in other respects. An account of exhaustiveness in *it*-clefts in terms of uniqueness presuppositions is very much in line with a prominent syntactic analysis of *it*-clefts as involving a covert definite determiner (Akmajian, 1970; Percus, 1997). On this account, then, the exhaustiveness effect is effectively reduced to a uniqueness/maximality presupposition triggered by the lexical meaning of the covert determiner (Frege, 1879; Heim, 1991). Conversely, a pragmatic account of exhaustiveness in *it*-clefts in terms of a generalized conversational implicature can explain the fact that exhaustiveness effects with *it*-clefts can sometimes be absent (Horn, 1981; Prince, 1978). On the pragmatic view, the observed exhaustiveness effect would result from default pragmatic strengthening procedures on par with the default exhaustive interpretation of answers to wh-questions (Schulz & van Rooij, 2006; Spector, 2007).

In this contribution, we will not take a direct stand on the precise nature of the exhaustiveness effects in *it*-clefts as presuppositional (4iii) or as (generalized) conversational implicatures (4iv), except for a few tentative remarks in the discussion. Instead, we pursue the more modest goal of presenting experimental results which show that the violation of exhaustiveness in German *it*-clefts differs from the violation of exhaustiveness in the German only-focus construction, which seems to constitute a genuine semantic violation in form of a logical contradiction. By contrast, we show that the exhaustiveness effect in German *it*-clefts is not truth-functional in nature, which is in agreement with, and hence in support of, the theoretical literature.

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
2.2. Eye-tracking research on focus and it-clefts

It is well-known that listeners detect focused information more quickly than non-focused information (Cutler & Fodor, 1979), and can also remember it better than non-focused information (Singer, 1976). Like other types of focus, it-clefts also facilitate processing in online sentence comprehension tasks. This fact about focus has been confirmed in numerous studies. In a probe recognition and naming task, Birch and Garnsey (1995) observed that clefted nouns are named faster and more accurately than non-clefted nouns, suggesting that it is easier to access the clefted material. Similarly, in an eye-tracking study on pronoun resolution, Foraker and McElree (2007) found that clefted noun phrases were more available as antecedents for anaphoric pronouns than their non-clefted counterparts.

In related work, Sedivy, Carlson, Tanenhaus, Spivey-Knowlton, and Eberhard (1994) explored contrastive focus in the auditory modality in conjunction with visual processing. When contrastive stress was placed on an object in the visual display (the small blue circle) eye movements to the object were faster. When the contrast member (the large blue circle) was referred to in a later utterance, eye movements to the contrast member were faster in the contrastive focus condition. Thus, the main finding was that reference resolution of the focused object and the contrast member were facilitated by the presence of contrastive stress.

Similarly, in an eye-movement study of reading, Carpenter and Just (1977) found that when one of two referents was focused by a cleft or pseudocleft, and an ambiguous pronoun was present in a subsequent sentence, readers tended to make a larger proportion of regressions to the focused antecedent. The main finding was that readers are less likely to fixate on a non-focused antecedent, and this finding extends to clefts as well.

One reason that the clefted noun might have greater accessibility in memory may be that the focused item achieves a higher activation in memory because the clefted item ends up being encoded more richly. In a clefted structure like It was John who bought a book, an exhaustiveness interpretation must be computed: no-one but John bought a book. Processing the clefted noun would at least involve generating the exhaustiveness interpretation and associating it with the noun; these are probably the underlying processing steps that result in a richer encoding of the clefted noun in memory. Indeed, recent work by Hofmeister (2009) shows that this extra processing increases the activation of the clefted noun in memory, which in turn results in easier retrieval of the noun at a subsequent stage.

In summary, the evidence shows that clefting a noun makes it more accessible in memory once the noun has been encoded. Interestingly, evidence also exists for increased encoding cost due to clefting; this comes from an eyetracking study conducted by Birch and Rayner (1997). They compared clefted sentences such as (6a) with a non-clefted sentence (6b).

(6) (a) It was the suburb that received the most damage from the ice storm. (b) Workers in the suburb hurried to restore power after the ice storm.

They found twice as many regressive eye movements (regressions, or leftward eye-movements while reading have been associated with increased processing effort; see, e.g., Frazier and Rayner (1982)) to the first three words of the sentence in the clefted condition compared to the non-clefted one. In addition, the probability of regressing out of the clefted word was higher (but the effect did not reach significance; in subsequent work they did find a significant effect, see Experiment 3B in Birch and Rayner (2010)). Finally, longer reading times were seen on the clefted versus non-clefted word. Birch and Rayner (1997) suggest that this is evidence for the encoding cost arising from the processing events triggered by a cleft structure (although in Birch and Rayner (2010) they found the opposite pattern, shorter first-pass reading times in the clefted versus non-clefted noun).

Thus, evidence exists for increased encoding cost of clefting—which is probably due to the cost of generating the exhaustiveness reading— and for greater accessibility of the clefted noun after it has been encoded. The presence of both effects—increased encoding cost followed by facilitation during subsequent processing—has been found in a self-paced reading study conducted by Hofmeister (2009). Further, a Hindi eyetracking study involving clefts (Vasishth et al., 2010) also showed similar results:
increased encoding cost in clefted nouns, expressed as higher regression probability (similar to the result in Experiment 3B of Birch and Rayner (2010)), and easier processing of the clefted noun in re-reading time.

For the purposes of this paper, the central insight to take away from the work on focus in the eyetracking paradigm is that focused constituents in general and it-clefts in particular involve some processing cost initially (e.g., the higher regression probability on clefts found in Birch and Rayner (2010) and Vasishth et al. (2010)), and in it-clefts this increased processing cost is probably due to the comprehension mechanisms computing the exhaustiveness interpretation as soon as the clefted noun is processed.

2.3. Event-related brain potential studies on focus-marking

Event-related potentials (ERPs) provide an excellent means for investigating language processing online because ERPs have a very high temporal resolution (Kutas and van Petten (1994)). ERP effects (components) are characterized by a set of quantitative (peak latency) and qualitative parameters (polarity, topography, experimental sensitivity). Distinct ERP patterns have been found in response to linguistically distinct experimental manipulations. They are identified by a nomenclature which refers to their polarity (N/negativity versus P/positivity), post-stimulus peak latency and topographic distribution. Four main markers of language processing have been identified in the literature: the early left anterior negativity (ELAN), the left anterior negativity (LAN), the N400 and the P600. However, in the following we will only discuss the N400 and the P600 component because of their relevance for the interpretation of our experimental findings.

**N400**: The N400 is a negativity with a latency peaking typically around 400 ms after the onset of a critical element. It has a centro-parietal bilateral (sometimes also global) distribution often with a slight right hemisphere focus. It has been shown that this component reflects the cost of (semantic) integration of non-stereotypical events and/or elements which do not fit with the (extra) linguistic context-expectancy (e.g., Kutas & Hillyard, 1980; 1983; Nieuwland & Kuperberg, 2008). Compare (7a) and (7b).

(7) (a) He drank a glass of coke.
(b) He drank a glass of rocks.

Additionally, Kutas and Hillyard (1984) and Kutas and Federmeier (2000) showed that the N400 is modulated by the structure of semantic memory. In other studies, N400 effects were elicited by lexico-semantic factors as well as thematic role assignment (e.g., Hoeks, Stowe, & Doedens, 2004; Friederici & Frisch, 2000). Semantic integration and expectancy N400s were also reported at discourse level (e.g., van Berkum, Hagoort, & Brown, 1999).

Crucially for the discussion to come, the N400 cannot be taken as an indicator of truth-value violations. Fischler, Bloom Childers, Roucos, and Perry (1983) showed that sentences that differ in truth value but that contain semantically related words do not generate an N400. For instance, no differences in N400 amplitude were found for the false (8a) versus the correct (8b). Compare this with the N400 effects in (9a) versus (9b).

(8) (a) #A robin is not a bird.
(b) A robin is a bird.

→ no difference in the N400 amplitude

(9) (a) A robin is not a tree.
(b) #A robin is not a bird.

→ an enlarged N400 amplitude for the false and non-fitting (a) compared to (b)
The so-called P600 (late Positivity) is a positivity peaking between 600 and 900 ms with a centro-parietal distribution and has been associated with (syntactic) reanalysis and repair (e.g. Osterhout & Holcomb, 1992). Additionally, this component has been found in response to enhanced syntactic complexity (e.g. Kaan, Harris, Gibson, & Holcomb, 2000; Friederici, Hahne, & Saddy, 2002). However, in recent work the idea of a purely syntactically related component has been questioned. Some studies showed the P600 component can also be induced when a semantic anomaly or violation (2.3) was detected (e.g. van Herten, Kolk, & Chwilla, 2005); see also Kolk, Chwilla, van Herten, and Oor (2003). However, the interpretation of a semantic P600 is not without controversy—for example, Bornkessel-Schlesewsky and Schlesewsky (2008) argued that the phenomenon of semantic P600s is language-specific and is only found in languages with fixed word order (English and Dutch). Similar manipulations elicit N400 effects in languages with a free word order (e.g. German).

(10) *De kat die voor de muizen vluchtte rende door de kamer
‘The cat that fled from the mice ran across the room (paraphrase)’

Additionally, a P600 effect/Positivity was found in licensing violations of German negative polarity items in (11a) versus (11b) (Drenhaus, Saddy, & Frisch, 2005; 2006). The (semantic/pragmatic) P600 follows directly under the assumption of semantic/pragmatic licensing of negative polarity items (e.g. Ladusaw, 1980).

(11)(a) Kein Mann, der einen Bart hatte, war jemals glücklich.
‘No man who had a beard was ever happy’

(b) *Ein Mann, der keinen Bart hatte, war jemals glücklich.
‘A man who had a beard was ever happy’

→ P600 on the unlicensed negative polarity jemals ‘ever’ in (11b) compared to the grammatical licensing environment (11a).

Since a logical contradiction constitutes a semantic violation one might therefore expect a P600/Positivity to show with exhaustiveness violations with only-foci and it-clefts if the exhausted meaning component were to constitute a conventionalized part of the meaning of these structures.

Focus structure, it-clefts and the violation of information structure: With regard to the topic of our paper, we will highlight only some of the studies which are relevant for the discussion of our data later.

Cowles, Kluender, Kutas, and Polinsky (2007) tested in an ERP study whether readers made use of focus structure to resolve the information structure of sentences (see also, Cowles, 2003). They presented participants with contexts followed by congruent target sentences (12a) or incongruent target sentences (12b), respectively.

(12) Context: A queen, an advisor, and a banker were arguing over taxes. Who did the queen silence with a word, the banker or the advisor?
(a) It was the banker that the queen silenced.
(b) It was the queen that silenced the banker.

They found a lateralized negativity between 200 ms and 500 ms on the noun queen (12b) compared to banker (12a) which the authors interpreted as an N400. Additionally, the authors reported a large positivity (P3b) for all words in the focus-marking (cleft) position. This positivity was interpreted as a marker for integration which was triggered by the focused elements in the question-context. In general, the authors highlight with their results that readers are using restrictions of the discourse and structural focus information online to process sentences.

Bornkessel, Schlesewsky, and Friederici (2003) tested the processing of clause-medial word order variations in German question answer-pairs. They observed a negativity for object-initial sentences in
a neutral context which was explained as a reflex of local processing difficulty. Additionally, they found a parietal positivity (280–480ms) on the focused (questioned) initial objects and initial subjects, respectively. Bornkessel et al. (2003) suggested that the elicited positivity can be interpreted as an effect of structural focus processing or rather focus integration.

Stolterfoht and Bader (2004) studied word order variations in German. The structures in (13) are ambiguous. Until the parser reaches the auxiliary, the subject of the canonical structure (13a) or of the non-canonical structure (13b) can either be die neue Lehrerin ‘the new teacher’ or einige der Kollegen ‘some the colleagues’. Additionally, the sentences in (13) have a different focus structural representation. (13a) has a wide focus reading, whereas a narrow focus on the second NP and the verb is required for (13b). In other words, the disambiguation in structures like (13b) would not only involve a reanalysis process but also an adjustment of the focus structure (under the assumption that the focus structure in (13a) is the default structure [Bader, 1998]).

(13) (a) ..., das [die neue Lehrerin\textsubscript{subject} einige der Kollegen\textsubscript{object} angerufen hat\textsc{FOC}.
... that the new teacher some the colleagues phoned has.
... that the new teacher phoned some of the colleagues.

(b) ..., das die neue Lehrerin\textsubscript{object} einige der Kollegen\textsubscript{subject} angerufen haben\textsc{FOC}.
... that the new teacher some the colleagues phoned have.
... that some of the colleagues phoned the new teacher.

The authors found a positivity and a right central negativity (500–600ms) for structures like (13b) at the disambiguation point (auxiliary). They interpreted the positivity as the reflex of syntactic reanalysis and the negativity as a focus structural revision process. Finally, Stolterfoht, Friederici, Alter, and Steube (2007) investigated the processing of contrastive ellipsis (replacives) with the focus particle nur ‘only’ in German. In their ERP study, they found on the contrasted constituent a positivity (350–1300ms) and a negativity between 450 ms and 650 ms. The positivity was interpreted as a correlate of focus structural processing and the negativity as a reflection of implicit prosodic processing.

The above overview of focus, it-clefts and the violation of information structure does not give a conclusive picture of the interpretive properties of the focus-clefted constituent. While some studies reported positivities; others found negativities. The reason for these differences might be that these studies tested and compared different types of structures (e.g. complement clauses versus it-clefts). Additionally, some studies examined sentences with context and other studies tested sentences without context.

2.4. Predictions

In order to test for the nature of the exhaustiveness effect with it-clefts and only-foci in German, we compare the ERP patterns in the processing of (marked) non-exhaustive German it-clefts (14b) versus their (unmarked) exhaustive counterparts (14a); we also compare the processing of (marked) violations of exhaustiveness in German only-foci (15b) versus the processing of their (unmarked) exhaustive counterparts(15a).

(14) (a) Es ist Maria, die das Klavier spielen kann und außerdem noch die Geige, sagte...
\textit{[it -cleft, [+exh]: unmarked]}

(b) Es ist Maria, die das Klavier spielen kann und außerdem noch Luise und Jana, sagte...
\textit{[it -cleft, [-exh]: marked]}
It is Mary that plays the piano and, besides, the violin [+exh]/Luise and Jana [- exh], said ...

(15) (a) Nur Maria kann das Klavier spielen und außerdem noch die Geige, sagte...
\textit{[only -focus, [+exh]: unmarked]}

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
(b) Nur Maria kann das Klavier spielen und außerdem noch Luise und Jana, sagte… 
[only -focus, [-exh]: marked]
Only Mary can play the piano and, besides, the violin [+exh]/Luise and Jana [-exh], said …

If it-clefts and only-foci have an identical truth-functional interpretation, we would expect the (a) and (b) sentences in the pairs in (11) and (12) to pattern alike. If, on the other hand, the two structures were to involve different processes, we would expect distinct ERP patterns to arise.

3. Experimental investigations

3.1. Questionnaire

As a pretest we carried out a questionnaire study. The goal was to investigate the acceptability of violations of exhaustiveness in German it-clefts and only-foci structures (14 b and 15 b) compared to their unmarked counterparts (14 a and 15 a).

3.1.1. Participants

80 students of the university of Potsdam (mean age 23 years, 17 male) participated in this experiment after giving informed consent. They were all monolingual speakers of German and received course credits for their participation. 14 participants were excluded from the analysis because they did not fill in the questionnaire, or because they judged semantically or syntactically ungrammatical filler sentences as grammatical.

3.1.2. Methods and design

Participants rated the acceptability of 40 sentences on a six point scale, with 1 representing the highest and 6 representing the lowest degree of acceptability (German school marking system). Participants were instructed to judge the acceptability of the sentences. The extreme values of the scale were illustrated in the instruction section of the questionnaire. We designed 32 sets of critical sentences (it-clefts and only-foci). The sentences were split into 8 different versions of the questionnaire. Each version contained 4 marked and 4 unmarked it-clefts (14a and 14b) and 4 marked and 4 unmarked only-foci (15a and 15b). Additionally, there were 24 unrelated filler sentences (50% grammatical and 50% ungrammatical).

3.1.3. Results

On the scale from 1 (best) to 6 (worst) the marked (non-exhaustive) it-cleft condition got a mean rating of 2.8 (se 0.0756) compared to their unmarked (exhaustive) counterparts 2.4 (se 0.0723). In the only-foci condition we find a mean rating score of 3.7 (se 0.0844) for the marked (non-exhaustive) condition and a mean rating score of 2.6 (se 0.0797) for unmarked (exhaustive) condition, respectively (cf. Fig. 1).

The statistical analysis on participants’ ratings was done by fitting a linear mixed model (LMM) Bates and Sarkar (2007), We used participants and items as random factors. The exhaustiveness violation in it-clefts and in only-foci were used as fixed factor (contrast coding in each case: −0.5 for the (marked) non-exhaustive condition, and +0.5 for the (unmarked) exhaustive condition). Additionally, we calculated the the interaction in order to test whether the exhaustiveness violation has a different impact in it-clefts compared to only-foci. In LMMs, the degrees of freedom are difficult to compute due to partial pooling; however, an absolute t-value of 2 or greater corresponds to significance at the a level 0.05 (Gelman & Hill, 2007).

The analysis revealed that there is a significant effect for the exhaustiveness violation in it-clefts (coefficient 0.3552, se 0.0938, \( t = 3.786 \)) and for the exhaustiveness violation in only-foci (coefficient 1.0543, se 0.1299, \( t = 8.115 \)). Additionally, we found an interaction (coefficient −0.34937, se 0.06596, \( t = −5.297 \)). The results show that, on the one hand, participants rated an exhaustiveness violation less acceptable in it-clefts and in only-foci respectively. On the other hand, the interaction suggest that the violation of exhaustiveness in only-foci is less acceptable than in it-clefts. In other words,
exhaustiveness or rather the violation of exhaustiveness may have a different status in the tested structures\footnote{An anonymous reviewer points out that the only-violation received only a mean rating of 3.7 compared to the mean rating of 2.8 in it-clefts violations. It may be surprising to the reader that the ratings do not have more extreme values. It is possible that the ungrammatical filler sentences (violation of negative polarity licensing, violation of the word order in subordinated sentences as well as phrase structure violations) had an impact on the judgment in the questionnaire; the only-violation and the it-cleft violation may have received a better score had these fillers not been present. However, the important point here is that we are interested in relative differences between the two sets of conditions; absolute scores are not central to the discussion.}

3.2. The ERP study on it-clefts and only-foci

3.2.1. Participants
16 undergraduate students (mean age 24 years, 2 male) from the University of Potsdam participated in this experiment towards fulfilling curricular requirements; participants gave informed consent. All participants were monolingual, right-handed and had normal or corrected-to-normal vision.

3.2.2. Material
We constructed 4 versions with 92 critical sentences each (23 per condition) like those in (7) and (8) intermixed with 92 unrelated filler sentences from other experiments.

3.2.3. Procedure
Participants were seated around 100 cm in front of a 17" computer screen in a separately grounded SIEMENS pure-iron shielding cabin. Sentences were presented in grey letters on a blue background in a 28-point Times font. Each sentence was preceded by a fixation cross. The sentences were presented in a pseudorandomized order. EEG recording was performed on an IBM-compatible Pentium I using the recording software Xrefa (Novagk, 1998). ERTS (Experimental Run Time System, Beringer (1996)) was used to present the stimuli and to mark the relevant points (setting triggers). ERTS was installed on an IBM-compatible Pentium PC connected to the EEG recording computer. Responses were recorded with

Fig. 1. Questionnaire: it-clefts and only-foci (Means with confidence intervals (95%))

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
the ERTS ExKey Keyboard Logic System (BeriSoft Corporation, Beringer, 1996). First, 12 training sentences were presented. After this training set, the 92 critical sentences and the 92 filler sentences were randomly presented in the center of a screen, with 400 ms (plus 100 ms interstimulus interval) for each unit (indicated by brackets in (16)).

(16) **it-cleft [+exh]**: [Es ist Maria,] [die] [das Klavier] [spielen kann] und [außerdem noch] [die Geige,] [sagte] [der Lehrer.]

**it-cleft [-exh]**: [Es ist Maria,] [die] [das Klavier] [spielen kann] und [außerdem noch] [Luise und Jana,] [sagte] [der Lehrer.]

**only-focus [+exh]**: [Nur Maria] [kann] [das Klavier] [spielen] [und] [außerdem noch] [die Geige,] [sagte] [der Lehrer.]

**only-focus [-exh]**: [Nur Maria] [kann] [das Klavier] [spielen] [und] [außerdem noch] [Luise und Jana,] [sagte] [der Lehrer.]

500 ms after the last phrase of each sentence a single noun was presented on the screen for 400 ms (Probe detection). The task was to judge within a maximal interval of 3000 ms by pressing one of two buttons whether this noun was part of the sentence the participant had read before (probe detection). The distribution of probes was balanced over the conditions (50% correct probes and 50% incorrect probes, respectively). 1000 ms after their response, the next trial began. The EEG was recorded by means of 16 Ag/AgCl shielded electrodes fixed in elastic caps (EasyCap) with a sampling rate of 250 Hz (with impedances < 5 kOhm) and were referenced to the left mastoid (re-referenced to linked mastoids offline). Following the nomenclature proposed by the American Electroencephalographic Society (Sharbrough et al., 1995) the electrodes were placed on the scalp. The horizontal electrooculogram (EOG) was monitored with two electrodes placed at the outer canthus of each eye and the vertical EOG with two electrodes above and below the right eye. The data from all channels was recorded with a 32-channel PORTI-32/MREFA TMS-Amplifier. During recording no online filters were used.

3.3. Behavioral data—probe detection

Accuracy percentages and response latencies of the probe detection task (cf., Table 1 and Figures 2, 3) were analyzed, even though there was no specific hypothesis for this data. It is important to mention that the probe detection task was a measure that was intended to ensure that participants really attended to words in the experimental sentences.

We performed a statistical analysis for accuracy and log-transformed reaction times by fitting a linear mixed model (LMM) Bates and Sarkar (2007). For analyzing accuracy we used a generalized linear mixed model with a logistic link function. The violation of exhaustiveness in **it-clefts** was used as fixed factor (contrast coding: −0.5 for the (marked) non-exhaustive condition, and +0.5 for the (unmarked) exhaustive condition) and **only-focus**-conditions (contrast coding: −0.5 for the (marked) non-exhaustive condition, and +0.5 for the (unmarked) exhaustive condition). Participants and items were used as random factors. In the **it-clefts**, no effect was found in the accuracy measure of (marked) exhaustiveness violation. However, there was an effect in the reaction time data (coefficient 0.0693, se 0.0255, t = 2.71). In the **only-foci** analysis we found an effect for (marked) exhaustiveness in accuracies (coefficient −1.2365, se 0.3826, z = −3.322) and reaction times (coefficient 0.0687, se 0.0254, t = 2.72).

An anonymous reviewer pointed out that in these types of tasks participants might not read normally for comprehension, but instead build a set of possible probes. It might be that probe detection induces a strategy of some sort. However, the results are replicated across methods. Our results are in line with the results of an offline study by Onea and Beaver (2009). Hence, we think that participants are not just doing probe monitoring in our experiment. Crucially, if participants were following probe-monitoring strategies it is is hard to explain why we find different pattern for **it-clefts** and **only-focus** in our ERP-results.
3.4. The ERP data

Only trials without artifacts were selected for the ERP analysis. The ERP data were filtered with 0.2 Hz (high pass) to compensate for drifts. Single-participant averages were computed in a 1200 ms window relative to the onset of the critical item (die Geige ‘the violin’ and Luise und Jana ‘Luise and Jana’) and aligned to a 200 ms pre-stimulus baseline. Two time windows were analyzed: 400–600ms for the N400 and 600–800ms for the P600 effects. Additionally, we conducted a third analysis (300–800ms) in order to investigate the differences between it-clefts and only-foci in the same time window.

3.4.1. Descriptive results

The ERP patterns from the onset of the critical element (die Geige ‘the violin’, onset at 0 ms up to 1200 ms thereafter) are displayed in Figs. 4 and 5. As can be seen from Fig. 3 (it-cleft conditions), ERPs in the (marked) non-exhaustive (14b) condition show a negativity that is broadly distributed compared to the (unmarked) exhaustive condition (14a). The negativity starts at around 400 ms and has a time window between 400 and 600 ms. From the ERP literature this effect can be interpreted as an N400 effect (see, section 2.3 and section 3.4.2 for the statistical analysis). By contrast, the ERPs in the only-foci conditions reveal a centro-parietal distributed positivity at between 600 and 800 ms for the (marked) non-exhaustive (15b) condition compared to the (unmarked) exhaustive condition (15a), see Fig. 4. This positivity can be interpreted as an P600 effect (see, section 2.3 and section 3.4.2 for the statistical analysis).

3.4.2. Statistical analyses

We fit a linear mixed model (LMM) Bates and Sarkar (2007) with ERP values averaged over items for each participant as dependent measure. Fixed factors were the exhaustiveness violation in clefts
(contrast coding: −0.5 for the non-exhaustive condition, and +0.5 for the exhaustive condition) and only-focus conditions (contrast coding: −0.5 for the non-exhaustive condition, and +0.5 for the exhaustive condition). Participants were used as a random factor.

As mentioned above, in the *it*-cleft conditions, visual inspection revealed a globally distributed N400 pattern in the 400–600 ms window. The LMM analysis, collapsing over the nine channels (F3, FZ, F4, C3, CZ, C4, P3, PZ, P4), showed that this effect was significant (coefficient −1.9544, se 0.6964, \( t = −2.8 \)). The centro-posterior positivity in the 600–800 ms window did not reach significance (\( t < 1 \)). In the only-focus conditions, the observed centro-parietal positivity in the 600–800 window was statistically significant (coefficient 4.5810, se 1.1131, \( t = 4.12 \)); this positivity was unaffected even when the three frontal electrodes were included in the analysis.

In order to compare the ERP response in *it*-clefts and only-foci within the same time window, we carried out a further analysis over a longer time window (300–800 ms); all nine electrodes were collapsed together. Here, we compared statistical models that either had only participants as random effects, or also had a separate slope term for each participant for the *it*-clefts and only-foci conditions (the so-called random intercepts and slopes model (Gelman & Hill, 2007)). The model with random slopes and intercepts was the best-fitting model, and showed no effect of exhaustiveness violations in clefting, whereas in the only-focus condition a positivity was found (*it*-clefts coefficient −0.6050, se 0.5781, \( t = −1.047 \); only coefficient 3.0463, se 0.6546, \( t = 4.654 \)).

4. Discussion

In the present study, we found a globally distributed negativity (400–600 ms window) when an exhaustiveness violation occurred in *it*-clefts, but a centro-parietal positivity (600–800 ms window) when an exhaustiveness violation occurred in only-foci. The negativity observed with *it*-clefts can be interpreted as an N400 effect (compare, Kutas & Hillyard, 1980; 1983; Nieuwland & Kuperberg, 2008). As mentioned in section 2.3, this effect can be taken as an indicator of a semantic/pragmatic integration process which is typically triggered by the processing of non-stereotypical events. In the case of non-exhaustive *it*-clefts, we contend that the non-stereotypical event is constituted by the violation of exhaustiveness itself. In this connection, it is important to keep in mind that truth-value violations do not generate an N400 effect (Fischler et al., 1983). This suggests that the violation of exhaustiveness with *it*-clefts does not constitute a semantic violation.

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
Bornkessel et al. (2003) and Cowles (2003); Cowles et al. (2007) reported increased positivities for clefts and interpret this as an attempt to integrate information. Cowles et al. (2007) observed an increased negativity (or less positivity) for an incongruent clefted noun compared to the congruent clefted noun. The authors interpreted their effect as an N400. They suggested two alternative ways to interpret their findings; namely, incongruence between focus assignment and preceding context or as a disruption of thematic processing. As they pointed out, both suggestions are compatible with a semantic interpretation of the N400 effect. However, none of the studies mentioned directly

---

**Fig. 4. It-clefts:** ERP effects on the critical element(s) die Geige 'the violin' and Luise und Jana 'Luise und Jana' from the onset at 0 ms up to 1200 ms at a subset of nine electrodes. Negativity is plotted upwards. The solid line displays the unmarked (exhaustive) condition and the dotted line displays the marked (non-exhaustive) condition. For presentation purpose only, ERPs were filtered offline with 8 Hz low pass.

Bornkessel et al. (2003) and Cowles (2003); Cowles et al. (2007) reported increased positivities for clefts and interpret this as an attempt to integrate information. Cowles et al. (2007) observed an increased negativity (or less positivity) for an incongruent clefted noun compared to the congruent clefted noun. The authors interpreted their effect as an N400. They suggested two alternative ways to interpret their findings; namely, incongruence between focus assignment and preceding context or as a disruption of thematic processing. As they pointed out, both suggestions are compatible with a semantic interpretation of the N400 effect. However, none of the studies mentioned directly
considers the consequences of violating the exhaustiveness interpretation generated by *it*-clefts and *only*; this makes it difficult to directly relate their findings to ours.

The negativity associated with *it*-clefts in our study can be interpreted as a reflex of integrating an unexpected element pragmatically (since truth-value violations do not generate an N400, Fischler et al. (1983)). Specifically, the mechanism of exhaustiveness cancelation is hypothesized to operate as follows: when the parser reaches the critical region of the exhaustiveness violation (*Luise und Jana*) it expects a different grammatical role/different animacy information (Bornkessel & Schlesewsky, 2006; Friederici & Frisch, 2000; Frisch & Schlesewsky, 2001, 2005; Hoeks et al., 2004), and therefore attempts to delete

---

**Fig. 5.** *Only-foci*: ERP effects on the critical element(s) die Geige ‘the violin’ and Luise und Jana from the onset at 0 ms up to 1200 ms at a subset of nine electrodes. Negativity is plotted upwards. The solid line displays the unmarked (exhaustive) condition and the dotted line displays the marked (non-exhaustive) condition. For presentation purpose only, ERPs were filtered offline with 8 Hz low pass.

Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneuroling.2010.10.004
the exhaustiveness interpretation generated when the cleft was processed (recall that the eyetracking
evidence discussed in Section 2.2 suggests that such an exhaustiveness interpretation is generated at the
cleft). This cancelation process is presumably reflected in the negativity seen in it-clefts.

There are other reasons to believe that the negativity we find is of a different nature than that
observed by Cowles et al. (2007). The negativity reported in our study shows a stronger peak compared
to the study of Cowles et al. (2007). Additionally, Cowles et al. negativity started earlier and lasted longer
(200–800ms) compared to our negativity (400–600ms). In addition, compared to our study, a smaller
amplitude was elicited in Cowles’ study; this was perhaps a result of the experimental design (context
sentence followed by the target sentence). Evidence for this comes from van Berkum et al. (1999) and
van Berkum, Zwitserlood, Hagoort, and Brown (2003), who showed that the amplitude of the N400 is
reduced when a context sentence precedes the critical element in a certain sentence context.

Regarding the only-foci exhaustiveness violation, if the underlying mechanism for exhaustiveness
cancelation were identical to that of it-clefts, we would have expected to see a negativity on the critical
element, analogous to the it-cleft case. However, a positivity was found with only-foci. This can be
interpreted either as an indicator of general reanalysis, or even as an indicator of a semantic violation
(van Herten et al., 2005, see section 2.3). The positivity might be related to the results of Stolterfoht et al. (2007) (compare section 2.3) who interpreted their positivity as a reanalysis process and the
mapping of different linguistic information (following the model of sentence processing by Friederici
et al., 2002).

There is another possible interpretation for the positivity found in our study. For example, Coulson,
King, and Kutas (1998) and Gunter, Stowe, and Mulder (1997) argued that a late positivity can be
related to another component, namely the P3b (Donchin & Coles, 1988). This component is sensitive to
expectancy. The less an event is expected, the more chance there is that the event is due to erroneous
processing. Hence, it is seen as an indication for the actualisation and updating of a mental model (see
also, Polich, 2007). In this sense, the findings for only-foci suggest that the underlying mechanism for
exhaustiveness is qualitatively different from it-clefts, and may require the processor to check the
veridicality of the exhaustiveness violation in only-foci.

Finally, the comparison of the observed patterns with exhaustiveness violations in it-clefts and
only-foci in the larger 300–800 ms window showed no effect in the it-cleft condition, but a positivity in
the only-focus condition. Taken together, these results suggest that the exhaustiveness violations in
it-clefts versus only-foci involve qualitatively different processing mechanisms and can hence be seen
as involving different generators. This is consistent with claims in the theoretical literature that
exhaustiveness violations in it-clefts are not on par with those only-foci. They may constitute either
a presupposition failure or a cancelation of a conversational implicature.

From a cross-linguistic perspective, these findings tie in with the results of an experimental study
on the well-known exhaustiveness effect in the Hungarian preverbal focus position. Based on an offline
text completion task, Onea and Beaver (2009) show that there is indeed an exhaustiveness effect
observed with preverbal foc in Hungarian. However, this exhaustiveness effect is not as strong as the
one observed with the truth-functional exclusive particle csak ‘only’, contra claims in the literature by
É. Kiss (1998; 1999), among others.

Turning back to the primary focus of our study, observe, again, that the choice between the two
remaining options for the source of exhaustiveness, namely presupposition and conversational
implicature, is not simply a choice between two otherwise equal semanto-pragmatic analyses. Rather,
it has far-reaching consequences for our conception of the syntax–semantics interface with it-clefts.
If the exhaustiveness effect indeed turned out to be presuppositional in nature, it would seem to be hard-
wired in the linguistic form of the cleft, possibly in form of a covert definite description (as argued in
Percus (1997), see section 2.1). On its alternative analysis as a (generalized) conversational implicature
(Horn, 1981), by contrast, the exhaustiveness effect would not be coded in the structure of the it-cleft.
Instead, it would be the result of a default pragmatic strengthening procedure that hearers apply,
presumably in order to justify the use of a non-canonical, and uneconomical, cleft-structure by the
speaker.

Given this central difference in terms of structural coding, the two analyses make a number of
different predictions that can be empirically evaluated. We are currently pursuing these questions in
ongoing work.
5. Conclusion

This study is, to our knowledge, the first online study that demonstrates the difference between exhaustiveness cancelation in it-clefts versus only-foci that has been proposed in the theoretical literature (e.g., Horn (1981)). The comparison of ERP patterns in these two conditions suggests that the exhaustiveness effect is triggered by different factors in the two structural environments. In particular, the differences found with only-foci and it-clefts suggest strongly that the exhaustiveness effect in it-clefts is not a truth-functional effect; this is consistent with proposals in the linguistic literature on the topic. It remains unclear whether exhaustiveness is a conventional meaning component of it-clefts, or whether its frequent occurrence with it-clefts is merely due to (default) pragmatic strengthening; this is an issue we intend to address in future research.

References


Please cite this article in press as: Drenhaus, H., et al., Exhaustiveness effects in clefts are not truth-functional, Journal of Neurolinguistics (2010), doi:10.1016/j.jneurolinguistics.2010.10.004


