Cleft Exhaustivity:
A Unified Approach To Inter-Speaker And Cross-Linguistic Variability

Dissertation von

Joseph P. De Veaugh-Geiss

eingereicht bei der Humanwissenschaftlichen Fakultät
der Universität Potsdam zur Erlangung des akademischen Grades
Doctor philosophiae (Dr. phil.)

Universität Potsdam
Department Linguistik

Eingereicht am 12. August 2019
Verteidigt am 18. Februar 2020

Gutachter:
Prof. Dr. Malte Zimmermann, Universität Potsdam
Dr. Jacopo Romoli, Ulster University

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-44642
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-446421

Dedication

This work is dedicated to my parents, who—despite my resistance at a young age—taught me the value of education.

In memory of Luis Vicente (1979–2018).

Acknowledgements

I am deeply thankful to my first advisor Malte Zimmermann, who has supported me in far too many ways to count since starting my Bachelor's at Universität Potsdam in 2008. As a mentor over the past eleven years he has provided me with intellectual stimulation and has encouraged me to venture out into the world, both as a linguist and as a traveller. His enthusiasm for semantics, African languages and culture, literature, among many other things, is inspiring, and I am honored to have had the opportunity to work alongside him in various capacities.

I am also profoundly inspired by my second advisor Shravan Vasishth, who never stops pushing himself and others to do better. It is because of him that I both fear and revere the art of making statistical inferences. His influence on my view of what makes for good science is significant and will continue to shape me for years to come.

Another big inspiration in my academic life is Edgar Onea, who never fails to amaze me by the breadth of his knowledge on any number of topics, whether it be presuppositions or programming in Python. His patience with me when I failed to grasp the intricacies of a topic is to be admired, and I will never forget the hours spent in his former office in Göttingen going over the various theoretical issues relevant to cleft exhaustivity. Thank you for bearing with me!

Luis Vicente was my first advisor for my Bachelor's thesis in Potsdam and one of the most lucid thinkers I have ever met. Luis taught me how to address complex issues by breaking down bigger problems into smaller ones. No matter how complex a topic was I always felt that I understood it after talking with him. The field of linguistics lost one of its brightest minds when he passed away at far too young an age last year. He made me a better thinker, and for this I will forever be thankful.

Going back much further in time, I would be remiss not to thank my high school principal, David Thaden, who showed me great patience and empathy despite my stubborn resistance to guidance at that age. As a young misfit, I often found myself seated before him in his office, sometimes under serious circumstances. Nevertheless, Mr. Thaden treated me as an adult, although I must have made that difficult for him at times. That he choose to believe in me in those crucial years is why I am now here writing these acknowledgments to a PhD dissertation. He passed away in May 2019, and I regret I never told him thank you in person.

There are so many people who have shaped me in my academic life that to list them all would prove difficult, but two stick out from the group. First, Marta Wierzba: Marta has been a reliable friend since we started our studies together, and her calm, thoughtful approach to not only linguistics but life in general has on numerous occasions helped me move forward to face the next challenge. Second, Agata Renans: there are so many wonderful things I could say about Agata that to write them all here would require a dissertation-length text. I will simply say thank you for all of the laughs and the ways you showed me how beautiful formal semantics can be. I look forward to starting the collaborations we had to put on the back-burner until I submit this dissertation.

I am also very lucky to have had Emilie Destruel, Swantje Tönnis, and Nadine Bade as co-authors over the past five years. I thoroughly enjoyed working out the details of our ideas together, and I am thankful for the trust and support you showed me as we committed those ideas to paper. A non-exhaustive list of colleagues I would additionally like to thank include the following (in no particular order): Mira Grubic, Anne Mucha, Jacopo Romoli, Carla Bombi, Radek Šimík, Henry Fominyam, Mareike Philipp, Johannes Schneider, Clare Patterson, Clara Huttenlauch, João Verissimo, Lena Jäger, Daniela Mertzen, Titus von der Malsburg, Daniel Schad, Laine Stranahan, Dario Paape, all of the members of Shravan Vasishth's lab, David Beaver, Gisbert Fanselow, Jordan Chark, Andreas Walker, Nadine Theiler, Reginald Akuoko Duah, Audrey Bürki, and Reinhold Kliegl. For each chapter I have also included the publication-specific thank-yous as they appeared in the original.

As a card-carrying member of the Free Software Foundation Europe (FSFE), I would like to thank the following free and open source software projects which have provided the technical means to realize many of my goals as a researcher: Debian and other GNU/Linux operating systems, TUXEDO Computers, R Studio and R (plus numerous packages, in particular *lmer*, *brms*, *tidyr*, *ggplot2*), Stan, LaTeX and TexWorks (and various packages, including *linguex*), LibreOffice, OwnCloud/NextCloud and SyncThing, OpenSesame, Python/PyGame, OnExp, FireFox/Tor Browser Bundle and Thunderbird (plus various privacy-enhancing add-ons), Moodle, GitLab, GNU Mailman, Audacity, GIMP, Calibre, Okular, SearX, ChatB, and Jitsi. My work as a researcher would be much harder without these tools.

Finally, there are several people who have made my job possible by taking care of the minutiae of academic life. In particular, I would like to thank Elke Pigorsch, Anja Gollrad, Fabienne Salfner, Ines Mauer, Annett Eßlinger, Michaela Schmitz, Anikó Grellmann, Ralf Ditsch, and Claudia Rößling.

This research was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 317633480—SFB 1287, Project C02 "Limits of Variability in Interpretation", as well as the "Exhaustivity in *it*-Clefts" (ExCl) project in the XPrag.de priority program (SPP 1727).

Contents

Introduction

This thesis is a contribution to the growing body of experimental work on focus-background *it*-clefts (or comparable clefts cross-linguistically), specifically in terms of the exhaustive inference they give rise to, illustrated in example (1) below. Although English *it*-clefts will be used in this chapter for the example sentences, the clefts of interest also include German *es*-clefts, French *c'est*-clefts, and Akan *nà*-clefts; in order to remain underspecified with respect to the language under consideration, I will henceforth use the generic term clefts.

(1) It is Max who mixed a cocktail. (*cleft*)

⇝ Ben, Jens, Tom, ... did not mix a cocktail. That is, nobody other than Max mixed a cocktail. (*exh. inference*)

How to model exhaustivity in clefts is a hotly debated topic in the literature and one that has received much attention in recent years. Generally speaking, there are two main camps in the debate: there are (i) those that take the exhaustive meaning to be conventionally-coded in the cleft structure itself, a hard-wired semantic inference which is predicted to arise robustly and systematically (among others, Atlas & Levinson 1981, Percus 1997, Büring & Križ 2013); and (ii) those that take exhaustivity to be a non-conventionally-coded pragmatic inference and thus derived via other means, such as Gricean conversational maxims or discourse-pragmatics, and hence potentially defeasible in context (among others, Horn 1981, De Veaugh-Geiss et al. 2015, Pollard & Yasavul 2016, Destruel & De Veaugh-Geiss 2019). Moreover, there are cross-linguistic differences reported in the literature in terms of the exhaustive interpretation of clefts, and given such differences some authors have proposed that the exhaustive inference is conventionally-coded for some languages, but not for others (e.g., the Kwa languages Akan and Ga vs. the West Chadic language Ngamo, discussed in Grubic, Renans & Duah 2019). Furthermore, the variability described above is reflected both in (disputed) introspective data as well as the results of recent experimental work. The four papers here on German, English, French, and Akan are a modest contribution to the debate.

First, some terminology. Focus-background clefts are a focus-partitioning device with the following form: the cleft *pivot* is the focused constituent appearing in the left periphery, which is followed by the cleft *predicate*, as shown in (2) (using the terminology in Križ 2017).[1] Preceding the cleft pivot there is often a pronoun plus copular verb (PRO + COP), although in some languages these are optional (see, for instance, Akan); whether this pronoun is expletive (Gazdar et al. 1985, Pollard & Sag 1994, É. Kiss 1999) or referential (Reeve 2007, 2012) is debated. In English, French, and German, the cleft predicate has the form of a relative clause headed by a relative pronoun (RELPRO);[2] by contrast, in a language such as Akan the pivot is partitioned from the predicate by a focus marker (FM) (although some authors have argued for a morphosyntactic and phonological relationship between the cleft predicate headed by the focus marker and restrictive relative clauses; see Titov 2019: §5 for discussion).

(2) (PRO + COP) *Cleft Pivot* FM/RELPRO *Cleft Predicate*

I will use a third term *referent*—adopted from Križ (2017)—in addition to the terms *pivot* and *predicate*. For the sake of concreteness, consider (3) below. The *referent* refers to the individual(s) who fulfill the cleft predicate in the discourse domain; thus, in (3) the *referent* is Max and Jens. By contrast, in the sentence in (3a), the *pivot* is Max.

(3) *Context*: Max and Jens mixed a cocktail. (*referent*: Max and Jens)
a. It is Max who mixed a cocktail. (*pivot*: Max)

1 Focus-background clefts are differentiated from topic-comment clefts in terms of whether the cleft pivot is the focus/topic and the cleft predicate the background/comment, illustrated in the examples (i)–(ii) from den Dikken (2013).

(i) Focus-background cleft
Q: What got you interested in clefts?
A: It was Brian's book that got me interested in clefts.

(ii) Topic-comment cleft
Q: Do you know Brian's book?
A: It was Brian's book that got me interested in clefts.

In the literature, focus-background clefts have been referred to as *stressed-focus* clefts (Prince 1978), *topic-clause* clefts (Hedberg 1990, Delin & Oberlander 1995), or *contrastive* clefts (den Dikken 2013), whereas topic-comment clefts have been referred to as *informative-presupposition* clefts (Prince 1978), *comment-clause* clefts (Hedberg 1990, Delin & Oberlander 1995), or *continuous-topic* clefts (den Dikken 2013). Jespersen (1927) coined the generic term cleft (Reeve 2007).

2 See, e.g., Percus 1997, Reeve 2007, 2012, and den Dikken 2013 for discussions about the syntactic structure of the cleft, which remains disputed.

$\rightsquigarrow$ Nobody other than Max mixed a cocktail. (*exh. violated*)

In cases when the referent properly contains the pivot (referent $\succ$ pivot), as is the case above, I will speak of a violation of the exhaustive inference.[3]

Important for the design of the studies here: similar exhaustive inferences are found across sentence types, such as in definite pseudoclefts (identity statements with a definite description), as in (4a); in sentences with an exclusive particle, as in (4b); and in canonical non-cleft constructions (for all languages considered in this thesis, this corresponds to subject-verb-object, or SVO, sentences), as in (4c).

(4) a. The one who mixed a cocktail is Max. (*pseudocleft*)
b. Only Max mixed a cocktail. (*exclusive*)
c. MAX mixed a cocktail. (*SVO*)

These sentence types play a crucial role in the studies reported in this dissertation. First, they serve as baselines for both asserted semantic exhaustivity (exclusives) and pragmatic exhaustivity (SVO). Furthermore, the sentences provide theoretically relevant points of comparisons to test various theories on the market, such as those that predict (i) direct parallels between clefts and definite pseudoclefts (among others, Boadi 1974, Percus 1997, Büring & Križ 2013, Križ 2016, 2017); or (ii) different response patterns in clefts vs. SVO sentences depending on predicate interpretation (see Renans & De Veaugh-Geiss 2019).

In all experiments presented here, the principal diagnostics used were judgment-tasks of exhaustivity violations, which follows in the footsteps of much of the experimental literature on exhaustivity inferences (e.g., Onea & Beaver 2009, Byram-Washburn, Kaiser & Zubizarreta 2013, Destruel et al. 2015). The idea behind such tasks is that violations of exhaustivity provide insight into the semantic or pragmatic source of the inference. Broadly speaking, if the exhaustive meaning is cancellable in context, then it must not be a conventionally-coded inference (asserted or presupposed), a claim that will be discussed further in this introduction; see also Chapter 1 (De Veaugh-Geiss et al. 2018), Chapter 2 (Destruel & De Veaugh-Geiss 2018), and Chapter 4 (De Veaugh-Geiss in preparation). In the experiments presented in Chapter 3 (Renans & De Veaugh-Geiss 2019), the authors used a slightly different type of violation, in which the distributive vs. non-distributive interpretation of the predicate may or may not—depending on the theory—give rise to a contradiction. This violation task allowed the authors to differentiate between various approaches to cleft exhaustivity. I will return to the above studies shortly.

The introduction will proceed as follows: The section "Research questions" outlines the general research questions and provides a brief outlook on how they

3 Although I am using the terminology in Križ 2017, I am not advocating for the analysis proposed there.

will be addressed in the four chapters of this dissertation. In the section "Exhaustivity inferences across sentence types", the main characteristics of clefts as well as definite pseudoclefts, exclusives, and SVO sentences are presented in terms of their exhaustive inference. This will be followed by the section "Characteristics of presuppositions", in which presuppositions (in general) and presupposition annulment (in particular) are discussed. The section "Summary of the conclusions" will provide an overview of the conclusions drawn across the four papers included here.

Research questions

In this section I outline the general research questions and four chapters of this dissertation. The experimental work presented here seeks to shed light on the following:

Semantic vs. pragmatic exhaustivity

- **Q1** When exhaustivity is violated, does cleft exhaustivity share characteristics with conventionally-coded *semantic* exhaustive inferences, such as assertions and presuppositions; or rather with non-conventionally-coded *pragmatic* exhaustive inferences, such as implicatures?
- **Q2** Do response patterns for clefts show parallels and/or differences to other sentence types which give rise to exhaustive inferences, namely definite pseudoclefts, exclusives, or canonical SVO sentences?

Modelling variability

- **Q3** How does one model the inter-speaker, cross-linguistic, and cross-predicate variability found in the data?

Throughout the four chapters of this dissertation, assorted types of variability are observed. First, there is the inter-speaker variation found in the different participant groups interpreting clefts—and definite pseudoclefts—either exhaustively or non-exhaustively, a variability in interpretation that reflects the ongoing debate in the diverse theoretical approaches found in the literature. Second, there are cross-linguistic differences, such as that found for French and English, which presents a unique challenge for a unified approach to cleft exhaustivity, since, as Destruel & De Veaugh-Geiss (2018: p. 13) [Chapter 2: p. 88] write, "whatever analysis one proposes for cleft's exhaustivity in one language must be able to explain the different data in the other". Finally, there is the cross-predicate variability in the acceptability of clefts in the series of sentences '*It is not* α *that did P.* α *and* β *did P.*', about which Velleman et al. (2012: p. 455) remark, "we must admit that [our intuitions] are subtle intuitions, and further empirical evidence here would be helpful". In Chapter

3, Renans & De Veaugh-Geiss (2019) in fact report that the variability of this series of sentences depends on the distributive vs. non-distributive interpretation of the predicate, as Velleman et al. (2012) suspected.

This dissertation seeks to address questions Q1–Q3 across four chapters consisting of two published papers (De Veaugh-Geiss et al. 2018 in Chapter 1, Destruel & De Veaugh-Geiss 2018 in Chapter 2), one paper accepted for publication (Renans & De Veaugh-Geiss 2019 in Chapter 3), and one paper in preparation (De Veaugh-Geiss in preparation in Chapter 4). For the published/accepted papers, I have faithfully reprinted them here, and any differences are purely presentational; note that when citing example, page, section, and other numbers from the published papers, I will refer to the numbers in the published papers, while also providing reference to the numbers in this dissertation in brackets when they differ. An overview of the four chapters is as follows:

Chapter 1

De Veaugh-Geiss, Joseph P., Swantje Tönnis, Edgar Onea & Malte Zimmermann. 2018. That's not quite it: An experimental investigation of (non-)exhaustivity in clefts. Semantics and Pragmatics 11(3). Early access.

DOI: https://dx.doi.org/10.3765/sp.11.3

In this paper, German *es*-clefts are compared experimentally to definite pseudoclefts, exclusives, and non-cleft SVO sentences using an incremental-retrieval information paradigm with a truth-value judgment task. By manipulating whether and when the exhaustivity inference is falsified, the authors are able to compare various theories of cleft exhaustivity, which they group into the (i) pragmatic (Horn 1981), (ii) semantic definite (Percus 1997, Büring & Križ 2013), and (iii) semantic Inquiry-Terminating (Velleman et al. 2012) accounts. The authors conclude that the semantic definite approaches to cleft exhaustivity—that is, the theories which predict parallels between clefts and definite pseudoclefts—were on the right track, but not quite right. Instead, they argue that the exhaustivity inference in both clefts and definite pseudoclefts is pragmatically derived from the anaphoric existence presupposition that is common to both constructions.

Chapter 2

Destruel, Emilie & Joseph P. De Veaugh-Geiss. 2018. On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts. Journal of Pragmatics. 138. 1–16. Second author was published under the name: DeVeaugh-Geiss, Joseph P.

DOI: https://dx.doi.org/10.1016/j.pragma.2018.09.009

In this paper, English *it*-clefts and French *c'est*-clefts are tested in comparison with exclusives and SVO sentences using a sentence-picture verification task. By employing a truth-value judgment task with response time measures, the authors explore both the interpretation and processing of clefts. Important differences have been observed between English and French, in particular in terms of the discourse-conditions in the use of clefts, making the two languages a relevant point of comparison for cleft exhaustivity. The authors propose that the discourse-pragmatic approach in De Veaugh-Geiss et al. 2018 can account for the cross-linguistic variability they report.

Chapter 3

Renans, Agata & Joseph P. De Veaugh-Geiss. 2019. Experimental Studies on it*-Clefts and Predicate Interpretation. Semantics and Pragmatics 12(11). Early access. Authors listed in reverse alphabetical order.*

DOI: https://dx.doi.org/10.3765/sp.12.11

In this paper, *it*-clefts and non-cleft SVO sentences in English are tested using an acceptability-judgment task in order to compare the *homogeneity* vs. the *alternative-based* approaches to cleft exhaustivity. By looking at how predicate interpretation—manipulated across contexts—affects the acceptability of the series of sentences '*It is not* α *that did P.* α *and* β *did P.*', the authors are able to differentiate between approaches which otherwise make the same predictions regarding the unacceptability of exhaustivity violations. The results of the three studies suggest that the acceptability of clefts does in fact depend on the interpretation of the predicate, thereby posing a challenge to the homogeneity approach of Büring & Križ 2013, Križ 2016, 2017.

Chapter 4

De Veaugh-Geiss, Joseph P. 2019. nà*-Cleft (non-)exhaustivity: Variability in Akan. Manuscript, Universität Potsdam.*

In this paper, the experimental methods used to test German *es*-clefts in De Veaugh-Geiss et al. 2018 are extended to Akan *na*-clefts and definite pseudoclefts, with exclusives and SVO sentences as controls. Despite the unforeseen response patterns in the incremental information-retrieval paradigm used in the studies, a post hoc exploratory analysis reveals by-participant variability comparable to that for German as reported in De Veaugh-Geiss et al. 2018 (Chapter 1). These results are compatible with a parallel approach to exhaustivity in both sentence types across both languages.

Exhaustivity inferences across sentence types

In this section, I will present an overview of some of the characteristics of exhaustivity inferences in clefts, definite pseudoclefts, sentences with exclusives, and non-cleft SVO sentences. The focus will be on the source of exhaustivity (assertion, presupposition, implicature), the at-issue vs. not-at-issue status of exhaustivity, projective behavior in embedded environments (typical for presuppositions), and cancellability in unembedded environments (typical for implicatures). The ambivalent conclusion I ultimately draw follows that of Büring & Križ (2013: p. 4), who write: "Our problem [. . .] is that exhaustivity is somehow implied in clefts, but it seems to be neither asserted, nor presupposed or implicated."

Clefts and definite pseudoclefts Focus-background clefts, as in (5), have been argued to give rise to at least three meaning components. There is (i) the asserted canonical meaning that is equivalent in meaning to the corresponding canonical SVO sentence, e.g., '*Max mixed a cocktail*'; moreover, there is (ii) the existence presupposition, e.g., '*Someone mixed a cocktail*'; and finally, there is (iii) the exhaustive inference, the central topic of this dissertation, repeated below for the reader.

(5) It is Max who mixed a cocktail.
  ⇝ Nobody other than Max mixed a cocktail. (*exh. inference*)

Definite pseudoclefts, i.e., identity statements with a definite description, are another sentence type of particular importance to this dissertation, shown in (6); similar to clefts, definite pseudoclefts are argued to give rise to an exhaustivity inference. In fact, in the literature definite pseudoclefts have been argued to be underlyingly parallel to clefts both syntactically and semantically (Akmajian 1970, Harries-Delisle 1978, Percus 1997, Hedberg 2000, Han & Hedberg 2008, Büring & Križ 2013, Križ 2016, 2017), and this has been argued to hold across multiple languages—of particular importance here, Akan (Boadi 1974, Ofori 2011, Titov 2019). Chapter 1 and Chapter 4 directly compare clefts to definite pseudoclefts experimentally, while Chapter 3 tests one of the theories—the homogeneity approach of Büring & Križ 2013, Križ 2016, 2017—which explicitly predicts parallels between the two sentence types.

(6) The one who mixed a cocktail is Max.
  ⇝ Nobody other than Max mixed a cocktail. (*exh. inference*)

For the sake of example, one approach, the semantic definite account in Percus 1997 (following the terminology in De Veaugh-Geiss et al. 2018 in Chapter 1), takes *it*-clefts to be syntactically derived from an underlying definite pseudocleft, and

therefore the two constructions are predicted to share the same semantic properties. The derivation is roughly as follows: from the definite pseudocleft in (7a), the relative clause is extraposed into clause-final position, shown in (7b); finally, a spell out rule for the definite results in *it* appearing clause-initially, shown in (7c).[4] As Percus (1997: p. 333) writes: clefts and definite pseudoclefts "are identical in their properties, and any account that applies to one should apply to the other". In this approach, clefts inherit the uniqueness or maximality presupposition of the definite pseudocleft, and thus exhaustivity is predicted to be conventionally-coded in both constructions via a combination of the uniqueness/maximality presupposition and the asserted identity statement, e.g., '*the unique* α *that* P*s is Max*'.

(7) a. The one(s) who mixed a cocktail is Max.
b. The one(s) t_i is Max [who mixed a cocktail]$_i$.
c. It t_i is Max [who mixed a cocktail]$_i$.

Exclusives Sentences with the exclusive particle *only* are the go-to example for at-issue, asserted exhaustivity. Generally, assertions are the foregrounded, at-issue content proffered by an interlocutor to be evaluated for truth or falsity; by contrast, presuppositions are the backgrounded information against which foregrounded content is evaluated. Example (8) illustrates the two primary meaning components of sentences with exclusives: whereas the exhaustive inference is an at-issue assertion, shown in (8a), the prejacent (the canonical sentence minus the exclusive) is often taken to be presupposed,[5] shown in (8b).

(8) Only Max mixed a cocktail.
a. Nobody other than Max mixed a cocktail. (*at-issue exhaustivity*)
b. Max mixed a cocktail. (*presupposition*)

An at-issue inference is one that directly answers the Question Under Discussion (QUD) (Roberts 1996). That the exhaustive inference of exclusives is at-issue—whereas it is not-at-issue in clefts and definite pseudoclefts—is demonstrated in the following example: The question in (9) is directly answered by the response with the exclusive (A), but not by the response with the cleft (A') or definite pseudocleft (A"). That is, since exhaustivity is at-issue in the sentence with the exclusive, the

4 In Percus 1997, the description in the definite pseudocleft is represented as a definite determiner plus a null head; here I have represented it with the definite determiner plus a number-neutral form, i.e., *the one(s)*.

5 However, see van Rooij & Schulz 2005, Beaver & Clark 2008, Roberts 2011, among others, for discussion of whether the prejacent is presupposed (in the sense of Karttunen 1974, Stalnaker 1974, etc). Since debates about the source of the prejacent are not relevant for the purposes of this dissertation, I will ignore them here.

response in (A) is informative; by contrast, the canonical meaning in the cleft and definite pseudocleft is at-issue, but exhaustivity is not-at-issue, and thus (A')–(A") sound redundant (Halvorsen 1978, Horn 1981, Büring & Križ 2013).

(9) Q: I know Max mixed a cocktail, but who else mixed a cocktail?
A: Only Max mixed a cocktail! (*exclusive*)
A': #It is Max who mixed a cocktail! (*cleft*)
A": #The one who mixed a cocktail is Max! (*pseudocleft*)

Not-at-issue inferences generally include implicatures and presuppositions; here I will concentrate on the latter. An important characteristic that distinguishes assertions from presuppositions is that assertions are visible to entailment-cancelling operators such as negation, modals, etc., whereas presuppositions (typically) are not. Consider, for instance, embedding under negation: whereas the asserted exhaustivity is interpreted under negation in (10a), the prejacent in (10b) is not—i.e., it projects out from under the negative operator, and thus (10) (still) entails that Max mixed a cocktail.

(10) Not only Max mixed a cocktail.
a. It is not the case that only Max mixed a cocktail (i.e., somebody else mixed a cocktail). (*exhaustivity negated*)
b. Max mixed a cocktail. (*prejacent projects*)

When clefts appear under negation, the canonical meaning is negated, shown in (11a). And if exhaustivity in clefts were presuppositional, one would expect it to exhibit projective behavior; however, as both Büring & Križ (2013: p. 3) and Velleman et al. (2012: p. 455) point out, among others, the exhaustive inference of clefts does not project out from under negation. This is illustrated in (11b). Here I am assuming, as I have up to now, that the exhaustive presupposition is *Nobody other than* α *did P*.

(11) It wasn't Mary who laughed.
a. It is not the case that it was Mary who laughed (i.e., Mary did not laugh). (*canonical negated*)
b. $\not\rightsquigarrow$ Nobody other than Mary laughed. (*exhaustivity does not project*)
[based on ex. (10) in Velleman et al. 2012]

For instance, if the above exhaustive inference did project out from under negation, one would predict the sentence in (12) to be unacceptable, contrary to fact.

(12) It wasn't Mary who laughed; it was Bill.
[ex. (10) in Velleman et al. 2012]

A similar pattern is found for definite pseudoclefts, shown in (13)–(14). Here I follow Križ (2017: Fn. 8, ex. ic) and have written the definite pseudocleft with the semantically appropriate number-neutral form '*the one(s) who . . . was/were*', which Križ concedes "is perhaps acceptable only in print" (p. 7).[6]

(13) The one(s) who laughed wasn't/weren't Mary.
⇝ Nobody other than Mary laughed. (*exhaustivity does not project*)

(14) The one(s) who laughed wasn't/weren't Mary. It was Bill.

Given that the exhaustive inference does not exhibit projection behavior but rather disappears when embedded under negation and other scope-taking operators, one might claim that exhaustivity in clefts and definite pseudoclefts does not have a presuppositional source but rather a pragmatic one, which brings me to the next sentence type of particular importance here.

SVO sentences Non-cleft canonical sentences, as shown in (15), are another sentence type with an exhaustive inference; but in contrast to the asserted exhaustivity of exclusives, exhaustivity in SVO sentences is generally analyzed as a not-at-issue pragmatic implicature. To indicate focus, signalled via pitch accent in English, the focused grammatical subject is written in all caps.

(15) MAX mixed a cocktail.
a. Max mixed a cocktail. (*assertion*)
b. Nobody other than Max mixed a cocktail. (*implicature*)

The not-at-issue status of the exhaustivity inference of SVO sentences is illustrated in (16), in which the response in B is uninformative; cf. the examples in (9) above.

(16) A: I know Max mixed a cocktail, but who else mixed a cocktail?
B: #MAX mixed a cocktail!

In unembedded environments pragmatic inferences are typically cancellable, and this is the case for exhaustivity in non-cleft SVO sentences, shown in (17a); compare to (17b), in which cancelling the at-issue truth-conditional exhaustivity of the exclusive results in unacceptability.

(17) a. MAX mixed a cocktail. In fact, Max and JENS mixed a cocktail.
b. #Only Max mixed a cocktail. In fact, Max and JENS mixed a cocktail.

6 As Križ (2017: Fn. 8) discusses, given the singular morphology of the form '*the one . . . who*', there may be an additional uniqueness presupposition that I wish to avoid here, and due to issues with number agreement the form '*the ones who . . . is*' is ungrammatical.

Similarly, when presented with a context such as (18), in which there is a plurality of cocktail-mixers, the sentence in (18a) remains true (albeit potentially degraded given the context) but the sentence in (18b) is clearly false. (Examples (18a)–(18b) in contexts such as (18) are directly comparable to the exhaustivity violations in the experiments presented in Chapter 1, Chapter 2, and Chapter 4.)

(18) *Context*: Max and Jens mixed a cocktail.
   a. MAX mixed a cocktail. ⇒ True
   b. Only Max mixed a cocktail. ⇒ False

Cancellability in *unembedded* environments—as in (17a) and (18a) above—is a characteristic typical of implicatures, whereas presuppositions are not cancellable in unembedded environments (e.g., Abrusán 2016; see also the following section for presupposition cancellation in *embedded* environments). Therefore, if cleft exhaustivity can be shown to be cancellable in unembedded environments, this may constitute evidence that the exhaustive inference is pragmatic in nature. However, examples such as (19a) suggest that exhaustivity is not (easily) cancellable in clefts, and the same holds for definite pseudoclefts such as (19b).

(19) a. #It was a pizza that Mary ate; indeed, she ate a pizza and a calzone.
   b. #The thing(s) that Mary ate was/were a pizza; indeed, she ate a pizza and a calzone. [adapted from ex. (18d) in Horn 1981]

Interim summary In a nutshell, unlike the at-issue asserted exhaustivity of exclusives, exhaustivity in clefts and definite pseudoclefts is a not-at-issue inference. Not-at-issue inferences typically include presuppositions and implicatures. However, exhaustivity in clefts and definite pseudoclefts does not project from under negation, as one expects for a presupposition. Moreover, exhaustivity in clefts and definite pseudoclefts is not easily cancellable in unembedded environments, as one expects for an implicature. To repeat Büring & Križ (2013: p. 4) from the opening paragraph of this section: "Our problem [. . .] is that exhaustivity is somehow implied in clefts, but it seems to be neither asserted, nor presupposed or implicated."

Characteristics of presuppositions

As De Veaugh-Geiss et al. (2018) point out (see Footnote 7 in Chapter 1), the definite pseudocleft construction was included in that study—as well as the study in De Veaugh-Geiss in preparation (Chapter 4), planned concomitantly—under the premise that exhaustivity was presuppositional, as is (typically) claimed for plain

definite descriptions.[7] Nevertheless, in light of the results in De Veaugh-Geiss et al. 2018 (Chapter 1) the authors conclude that the assumption of a presuppositional source for exhaustivity must be mistaken for the unembedded definite pseudoclefts tested there (i.e., those headed by the compound definite *derjenige*). In order to dig deeper into this claim, I will begin this section (i) by briefly exploring the general characteristics of presuppositions before (ii) delving into more detail into the matter of presupposition annulment, which is a cover term for when a presupposition does not enter the global context. Such cases include presupposition cancellation (in embedded environments) and presupposition dismissal, amendment, or rejection (in unembedded ones, of particular importance here). Following this section on presuppositions, I will wrap up the general introduction with a summary of the conclusions of the four chapters of this dissertation.

General characteristics of presuppositions As Karttunen (1973: p. 170) puts it, presuppositions are "the set of assumptions that the speaker of the utterance thinks [he or she] shares with [his or her] intended audience", referred to as the common ground. In other words, presuppositions are the backgrounded information against which the asserted foregrounded content can be evaluated. Generally, presuppositions are taken to be conventionally-coded as part of the meaning of specific lexical (e.g., the change-of-state verb *stop*) or structural (e.g., the existential in clefts) items, referred to as presupposition triggers. The dominant semantic view models presuppositions as partial functions which are only defined for contexts in which the presupposition of the sentence is satisfied (Heim 1992, Chierchia & McConnell-Ginet 1996, Heim & Kratzer 1998; compare to Simons 2001, Abusch 2002, Schlenker 2008, Abrusán 2011, 2016).

A typical characteristic of presuppositions is projective behavior, or the ability to be inherited by the global context when embedded under various entailment-cancelling operators (as discussed briefly for (10) above). Consider the cleft in (1), repeated in (20). I will concentrate here on the assertion, shown in (20a), and the existential presupposition, shown in (20b).

(20) It is MAX who mixed a cocktail.

a. Max mixed a cocktail. (*assertion*)

7 See, however, Szabo 2000, 2004, Ludlow & Segal 2004, Abbott 2014 for counter-arguments to a presuppositional account of uniqueness/maximality in plain definites. See also Šimík & Demian 2019 for experimental results for definite descriptions which are remarkably similar to those reported in Chapters 1 and 4 for definite pseudoclefts. Although Šimík & Demian (2019) entertain the possibility of the uniqueness/maximality inference in plain definite descriptions having a "conversational source", the authors ultimately reject this idea given differences in their studies between definite descriptions and standard pragmatic inferences.

b. Someone mixed a cocktail. (*presupposition*)

To test for projective behavior, a number of standard embedding environments called the Family of Sentences are used (Langendoen & Savin 1971, Chierchia & McConnell-Ginet 1996; see Levinson 1983, Kadmon 2001 for limitations to this diagnostic). Whereas the assertion that Max mixed a cocktail is questioned in (21a), interpreted under the conditional in (21b), and negated in (21c)—that is, the assertion is not entailed by any of the sentences in (21)—the presupposition that someone mixed a cocktail projects out from under the scope of the operators. That is, the existential of the cleft is neither questioned nor conditionalized nor negated. Thus, given the projective behavior, one can conclude that the existential inference is presupposed.[8]

(21) a. Is it Max who mixed a cocktail?
b. If it is Max who mixed a cocktail, . . .
c. It isn't Max who mixed a cocktail.
⇝ someone mixed a cocktail (*existential projects*)

Sometimes a presupposition is not in the common ground, though, as might be the case if the sentences in (20)–(21) are uttered to an interlocutor who does not know that someone had mixed a cocktail. In such cases, there are several ways a discourse may proceed. For instance, if the presupposition is not inconsistent with the context, it can be accommodated.[9] Lewis (1979: p. 340) defines accommodation like so: "If at time t something is said that requires presupposition P to be acceptable, and if P is not presupposed just before t, then – *ceteris paribus* and within certain limits – presupposition P comes into existence at t." In this way, accommodation amends the context in order to satisfy the missing presupposition. Conceived of as a repair strategy in Lewis 1979, accommodation is a process that has been reported to occur frequently in natural discourse in various corpus studies (Poesio & Viera 1998, Fraurud 1990, Delin 1992, Spenader 2002).

Accommodation is just one option, however: a discourse participant can also challenge the presupposition with a phrase like "*Hey, wait a minute. I didn't know . . .*" (the *Hey, wait a minute* test in von Fintel 2003). Such a response is usually unacceptable when targeting the at-issue asserted information. For example, an interlocutor might utter (B) as an acceptable response to the unknown presupposition in (22), whereas (B') targeting the assertion is odd.

8 See, however, Büring (2011: §5) and Büring & Križ (2013: §6), who argue against a hard-coded existence inference; cf. discussion by Velleman et al. (2012: §4.3).

9 The term was coined in Lewis 1979; for early approaches to accommodation, see Stalnaker 1974, Karttunen 1974, Grice 1981.

(22) A: It is Max who mixed a cocktail.
B: Hey, wait a minute! I didn't know someone mixed a cocktail.
B': #Hey, wait a minute! I didn't know Max mixed a cocktail.

When an unsatisfied presupposition results in inconsistency with linguistic or contextual information, accommodation leads to a contradiction between the presupposition and the context. Such presupposition failures can be due to the linguistic content alone, as in (23) (referred to as necessary presupposition failures in Heim & Kratzer 1998: §4.4.4), and those that are contextual, as in (24). For an example of a linguistic presupposition failure, the unacceptability of (23) arises due to the speaker expressing both (i) explicit ignorance about a letter-reading event as well as (ii) presupposing that someone having read the letter is common ground; for an example of contextual presupposition failure, the global context in (24) is at odds with the presupposition that someone mixed a cocktail.

(23) I have no idea whether anyone read that letter. #But if it is Bill who read it, let's ask him to be discreet about the content.
[ex. (4) in Abusch 2010]

(24) *Context*: No one mixed a cocktail.
#If it is Max who mixed a cocktail, ...

Presupposition annulment There are various ways, however, for the presupposition not to enter the common ground. I will use the umbrella term presupposition annulment for such cases. These include presupposition cancellation,[10] dismissal, amendment, or rejection. I will consider each in turn, with a focus on embedded vs. unembedded environments.

Embedding environments such as under negation—often followed by explicit denial—are contexts in which a presupposition may be cancelled.[11] As Abrusán (2016) writes: "Presupposition cancellation under negation, at least when reinforced by explicit denial of the content of the presupposition, is not particularly sensitive to the identity of the presupposition trigger: the content of any presupposition can be denied this way." Consider an example such as (25).

(25) It isn't Max who mixed a cocktail, because no one mixed a cocktail.

By the phrase "identity of the presupposition trigger", Abrusán (2016) is referring to the hard/soft distinction for presupposition triggers. Since the distinction is only

10 See Abrusán 2016 for the difference between cancellation and suspension; I will conflate the two terms here.

11 I will only discuss embedding under negation here, but there are other embedding environments in which presuppositions can be cancelled or suspended; see, e.g., Abrusán 2016.

relevant for (non-)cancellability in embedded environments—but unembedded environments are of primary interest here—I will not discuss the hard/soft distinction further (see, e.g., Abusch 2002, 2010, Romoli 2012, Abrusán 2016).

Presupposition cancellation has been analyzed in various ways. In one proposal, there is a presupposition-cancelling operator referred to as metalinguistic negation (e.g., Horn 1989). In another, the presupposition is locally accommodated under, e.g., negation, in essence by adding the presupposition to the asserted content (e.g., Heim 1983). (See, for instance, Renans & De Veaugh-Geiss 2019: §4 in Chapter 3 as well as Abrusán 2016 for discussion; cf. approaches in Gazdar 1979, Soames 1979.)

By contrast, in unembedded environments presuppositions are typically not cancellable (Abrusán 2016), a crucial point for the conclusions made in this dissertation. For instance, the sentence in (26) results in a presupposition violation: the definite description *the king of France* presupposes that there is a king of France, which is at odds with the assertion that there is no king of France. As a point of comparison, pragmatic implicatures, such as the exhaustive inference in (17a), repeated in (27), are generally cancellable in unembedded environments (Grice 1975) (cf. Kuppevelt 1996, Mayol & Castroviejo 2013 for some limits to implicature cancellation).

(26) #The King of France ate the cake. In fact, there is no King of France.
[adapted from ex. (1c) in Abrusán 2016]

(27) MAX mixed a cocktail. In fact, Max and JENS mixed a cocktail.

Thus, at first glance the following generalization appears to hold: in embedded environments presuppositions can be cancelled (i.e., via metalinguistic negation or local accommodation), but in unembedded environments they cannot.

This is not the full story, however. Unembedded presuppositions may fail to enter the common ground in various other ways. For instance, it has been observed that a presupposition may be left unaccommodated when not supported by the context, which I will refer to as presupposition dismissal (Tiemann 2014, Tiemann et al. 2011; but see Bacovcin, Zehr & Schwarz 2018). Tiemann (2014) captures this with the maxim of interpretation *Minimize Accommodation*, which states: "Do not accommodate a presupposition unless missing accommodation will lead to uninterpretability of the assertion!" (see also Moulton's (2007) *Accommodate Conservatively*). De Veaugh-Geiss et al. (2018) discuss the option of *Minimize Accommodation* in light of their data in Footnote 14 in Chapter 1.[12]

12 I do not discuss here the idea of being 'squeamish' about assigning a truth-value to a sentence with a presupposition failure (e.g., truth-value gap account in Strawson 1950, 1964; see also discussions in Abrusán & Szendrői 2013 and Križ & Chemla 2015). In all of the experiments with a truth-value judgment task presented in this dissertation, participants were only given two options: judging the sentence as either 'true' or 'false'. Although I cannot rule out the possibility that a sentence might be judged 'true enough' if a hearer felt squeamish, in one study which included a 'can't say' option—

Furthermore, it has been proposed that presupposed information may be subject to a reinterpretation process called *amendment* (Strawson 1954), in which a definite description with a presupposition failure can be updated or "corrected" to fit the context (compare to accommodation, when the context is updated to fit the presupposition). Strawson (1954: p. 230) writes, "where the speaker's intended reference is pretty clear, [what we can do] is simply to amend [his or her] statement in accordance with [his or her] guessed intentions and assess the amended statement for truth or falsity". For the sake of example, consider the situation of a speaker referring to "the person drinking a martini" when, in fact, the person's glass contains water. A listener can nonetheless identify the referent and simply 'correct' the failed presupposition. As a result, Strawson (1954: p. 230) observes, "we are not awarding a truth value at all to the original statement". As Donnellan (1966: Fn. 10) describes it: "The cases Strawson has in mind are presumably not cases of slips of the tongue or the like; presumably they are cases in which a definite description is used because the speaker believes, though [the speaker] is mistaken, that [(s)he] is describing correctly what [(s)he] wants to refer to".[13]

Finally, there are cases which I call rejection, i.e., when a listener rejects an assertion with a presupposition failure that is contradictory to world knowledge or

argued to correspond to squeamishness for reference failure with a definite description—participants generally judged positive sentences with presupposition failures as 'false' (see Abrusán & Szendrői 2013, discussed in this section). That said, unlike definite descriptions with respect to the existential, clefts would arguably be a "class 1" trigger, in Tiemann's terms, with respect to exhaustivity. For this class of triggers, assigning a truth-value to the assertion does not hinge on the contribution of the presupposition, and thus the presupposition "will be ignored rather than [. . .] accommodate[d]" when faced with a presupposition failure (Tiemann 2014: p. 44). If squeamishness leads to a "true enough" response, this arguably would be a case comparable to dismissal for such class 1 triggers.

13 Although De Veaugh-Geiss et al. (2018) in Chapter 1, Destruel & De Veaugh-Geiss (2018) in Chapter 2, and De Veaugh-Geiss (in preparation) in Chapter 4 do not explore the possibility of amendment as discussed above, I will do so here for clefts. Consider the example in (i), in which the referent of the cleft is unproblematic to identify, just as in the *martini* example above: the listener can easily 'correct' the existential presupposition of (ia) to fit the context.

(i) *Context*: Max is mixing a (non-alcoholic) lemonade. No one else is mixing a drink.

a. Speaker: It is Max who is mixing a cocktail.
⇝ *existential failure*: someone is mixing a cocktail
⇝ *amended*: someone is mixing a lemonade

For the sake of comparison, consider (iia) in the context of (ii). In this context, the existential is satisfied, but the exhaustivity inference does not hold.

(ii) *Context*: Max and Jens are mixing a cocktail.

a. Speaker: It is Max who is mixing a cocktail.
⇝ *exhaustivity failure*: Max and nobody else is mixing a cocktail
⇝ *amended*: Max (and somebody else) is mixing a cocktail

context, discussed in detail below. Beyond judging the statement as not true, an interlocutor may evaluate the sentence as having 'stopped making sense' should the presupposition be implausible in context (Singh et al. 2016). In cases of rejection the sentence with the presupposition failure is evaluated as false or nonsensical.

What all of the above strategies—i.e., presupposition cancellation, dismissal, amendment, and rejection—have in common is that no interpretation of the original presupposition has been inherited by the global context. Cancellation is a phenomenon that occurs in embedded environments, whereas dismissal, amendment, and rejection occur in unembedded environments. Of the latter three cases, dismissal or amendment may result in the utterance with the unsatisfied presupposition as being judged true, while in the case of rejection the utterance ends up being evaluated as false or uninterpretable.

In the few experimental studies employing variations of the tasks which are comparable to the violation tasks in the studies reported here, results suggest participants will opt for rejection when asked to make a judgment of a sentence with a presupposition failure. For instance, Abrusán & Szendrői (2013) tested global presuppositon failures of the existence presupposition in definite descriptions, both in embedded and unembedded contexts; I will focus on the latter here, since those are the cases most similar to the experiments in this dissertation. Abrusán & Szendrői (2013) used general knowledge of the actual world to trigger the failure. Consider the following example: the definite description in (28) lacks a referent and thus the presupposition is in contradiction with world knowledge. (Abrusán & Szendrői (2013) report that the items used were tested for accessible, general knowledge in pilot studies.)

(28) The king of France is bald.
[ex. C0 from Table 1 in Abrusán & Szendrői 2013]

What Abrusán & Szendrői (2013) found is that in the unembedded environments participants generally judged the sentence as 'false' (from a minimum of 77.8% to a maximum of 92.4% 'false' judgments for the six unembedded conditions; see Abrusán & Szendrői 2013 for more information), despite the option to select 'can't say' in their experimental design.

Even though the statement in (iia) has an inference failure in terms of exhaustivity, there is a person the speaker is referring to, namely Max, and that person can easily be identified by the listener. In fact, that is exactly the situation that Strawson (1954) was trying to capture, and I see *a priori* no reason why such a strategy could not be employed by participants in the studies here. In fact, amendment as a strategy would be compatible with the response behavior for roughly half of the participant population reported in Chapters 1 and 4. Nevertheless, amendment is not a formalized interpretive strategy, and it is not clear why half of the population would amend the statement with the exhaustivity violation while the other half would not. I will leave this issue for future research.

In another study, Romoli & Schwarz (2015) used a Covered Box paradigm to test local accommodation of the presupposition trigger *stop* when embedded under negation in sentences such as in (29). The change-of-state verb *stop* in (29) triggers a global presupposition that Benjamin went to the movies before Wednesday, but in the visual context this presupposition is violated. Thus, in contrast to Abrusán & Szendrői 2013, in which the violation was due to general world knowledge, in Romoli & Schwarz 2015 the violation was made explicit by the context.

(29) *Description of visual context*: Benjamin went to the beach on Monday–Tuesday and to the movies on Wednesday–Friday.

Target: Benjamin didn't stop going to the movies on Wednesday.

[ex. (28b) in Romoli & Schwarz 2015]

Romoli & Schwarz (2015) find that a majority of participants (71%) rejected the target by choosing the only acceptable alternative available, the covered box, instead of accommodating the presupposition under the negative operator (local accommodation was the phenomenon they were interested in; cf. Chemla & Bott 2013).

Šimík & Demian (2019) similarly employed the Covered Box paradigm for their studies on two languages with and without articles, namely German and Russian. They were primarily interested in testing uniqueness/maximality with singular/plural definite descriptions across the two languages; however, Šimík & Demian additionally report the results of the filler trials looking at various not-at-issue meaning components. Of particular interest here are the filler items with presupposition violations. These included sentences with the additive 'also' (i.e., *auch* in German and *tože* in Russian) as well as the determiner *both*, a universal quantifier presupposing a restrictor with cardinality two (Schwarzschild 1996). For the sake of space, I will concentrate here on the violation of the additive condition, illustrated in (30), in which the visual context contradicts the presupposition triggered by *auch* in the target sentence.

(30) *Description of visual context*: Julia and Thomas are taking a walk in the snow. Thomas is the only one wearing a cap.

Target:
Als Julia und Thomas spazieren gingen, ist es deutlich kälter
when Julia and Thomas walk.INF went is.AUX it clearly colder
geworden. Thomas hat auch eine Mütze aufgesetzt.
become Thomas has also a cap put.on
'*When Julia and Thomas went for a walk, it turned quite cold. Thomas put on a cap, too.*' [ex. (24a) in Šimík & Demian 2019]

Just like the results of the presupposition violation tasks in Abrusán & Szendrői 2013 and Romoli & Schwarz 2015, most participants rejected the sentence by selecting the black box alternative (96% of the time in German and 86% of the time in Russian)[14]; comparable results were found for the determiner *both* when the visual context depicted a scene with more than two relevant entities (the covered box was selected over 94% of the time).[15] Given studies like the above, it indeed seems that if a presupposed inference is contradicted in context, the result is that the sentence with the violation will be rejected when participants are asked to provide a truth-value judgment.

To sum up this section: Presupposition cancellation is possible in embedded environments; moreover, in unembedded environments there were at least three other ways a presupposition may fail to enter the common ground. Hearers may dismiss the presupposition when the presupposition is not necessary for interpretation of the assertion. Moreover, in some cases it may be possible to 'correct' the statement with the presupposition failure in order for the sentence to fit the context, a reinterpretation process referred to as amendment. Crucially, for either dismissal or amendment, it is possible for the assertion to end up being evaluated as 'true' despite the presupposition failure. Finally, hearers may altogether reject a sentence with a presupposition failure. Indeed, in the few studies which had explicit presupposition violations with truth-value judgment tasks, participants were found to reject a sentence with a violated presupposition a majority of the time. I will now conclude this introduction by providing a brief overview of the results and conclusions reported in the papers here.

Summary of the conclusions

In this section I will synthesize the results and conclusions in the four papers. I will focus on the cross-linguistic and inter-speaker variability in Chapters 1, 2, and 4 as well as the predicate variability in Chapter 3.

14 See also the experimental work on accommodation of the additive presupposition in German in Grubic & Wierzba 2019. The authors show for similar targets that participants are very likely to accommodate that the other mentioned individual (Julia) has the relevant property (wearing a cap). It is worth pointing out that additives such as *auch* are "class 1" triggers in Tiemann's (2014) categorization—nevertheless, in both Grubic & Wierzba 2019 and Šimík & Demian (2019) participants do not leave the presupposition unaccommodated or dismiss it, and when confronted with a direct violation they reject the sentence. Compare to discussion in Footnote 12 above; see also the study reported in Bacovcin, Zehr & Schwarz 2018.

15 By contrast, when uniqueness/maximality was violated for definite descriptions in German, the sentences were rejected only 42% of time, results which are directly comparable to those for definite pseudoclefts reported in Chapter 1 and Chapter 4.

The stimuli used in the violation task studies in Chapter 1 for German (De Veaugh-Geiss et al. 2018), Chapter 2 for French and English (Destruel & De Veaugh-Geiss 2018), and Chapter 4 for Akan (De Veaugh-Geiss in preparation) all have the same general exhaustivity violation: the referent of the predicate properly contains the cleft pivot (referent ≻ pivot) such that there is a violation of the exhaustive inference, illustrated below.

(31) *Context*: Max and Jens mixed a cocktail. (*referent*: Max and Jens)

a. It is Max who mixed a cocktail. (*cleft pivot*: Max)
⇝ Nobody other than Max mixed a cocktail. (*exh. violated*)

The linking hypothesis of such violation tasks is as follows: if exhaustivity can be cancelled when the target sentence is unembedded (i.e., there is no entailment-cancelling operator), then the exhaustive inference must *not* be conventionally-coded as part of the meaning of the sentence. Therefore, the general predictions are such: a sentence such as (31a) in the context of (31) will be judged 'false' if exhaustivity is coded as part of the meaning of the cleft; on the other hand, the sentence will be judged 'true' if exhaustivity is not coded as part of the meaning of the structure (modulo noise).[16] Thus, such tasks are intended to distinguish conventionally-coded semantic inferences such as assertions and presuppositions from non-conventionally-coded pragmatic ones such as conversational implicatures.

What is reported in the studies in Chapters 1, 2, and 4, however, does not fit this binary view. In fact, what all of the above studies report is that—when averaging over all participants (I discuss by-participant results shortly)—clefts elicit mid-range response patterns in exhaustivity violation contexts (ca. 40–50% judgments corresponding to a non-exhaustive interpretation). Moreover, in the studies that included definite pseudoclefts (Chapters 1 and 4),[17] the response patterns were on a par with clefts. Thus, clefts (and definite pseudoclefts) differed both from the asserted exhaustivity of exclusives, judged as strictly exhaustive, as well as from the pragmatic exhaustivity of SVO sentences, judged as weakly exhaustive. The exception to this was French, as reported in Chapter 2, in which *c'est*-clefts elicited overall weakly exhaustive responses, i.e., a high proportion of 'true' judgments (ca.

16 This is a simplified picture, since the (not-)at-issue status may have hypothetically also played a role in the judgments made in the incremental information-retrieval paradigm used in Chapter 1 (De Veaugh-Geiss et al. 2018) and Chapter 4 (De Veaugh-Geiss in preparation), which was one of the motivations for that design. In particular, see the discussion of *not-immediate* inferences in De Veaugh-Geiss et al. 2018: §3.4 [Chapter 1, Section 1.3.4]; however, note that clefts did not pattern as one might expect for such (hypothetical) inferences, and thus I will not discuss not-immediate inferences in this introduction.

17 The study on French and English in Chapter 2 (Destruel & De Veaugh-Geiss 2018) tested clefts vs. exclusives and SVO sentences, and not definite pseudoclefts.

75%) despite the violation; furthermore, no statistical difference with SVO sentences was found, in contrast to the same task for English. In all of the above studies, the authors argue that a truth-conditional analysis of cleft (and definite pseudocleft) exhaustivity is not compatible with the data.

In addition to offline truth-value judgments in the above studies, in Chapter 2 Destruel & De Veaugh-Geiss (2018) analyze response times per truth-value judgment made, which further reveals cross-linguistic differences between English and French. For French, when cleft exhaustivity was violated, not only did participants treat clefts as weakly exhaustive, but no difference in response times was found; by contrast, for English the responses corresponding to an exhaustive interpretation had significantly *faster* response times than those corresponding to a non-exhaustive interpretation. Destruel & De Veaugh-Geiss (2018) argue that the data for French and English present several interesting puzzles. In the offline data, exhaustivity in clefts appears to be cancellable—at first blush this seems most compatible with a non-conventionally-coded pragmatic analysis. However, with regards to the online response time measures, the literature on scalar implicatures has found that the enriched pragmatic meaning generally elicits *slower* response times—or at least not faster—compared to the literal meaning (among others, Bott & Noveck 2004, Breheny, Katsos & Williams 2006, Huang & Snedeker 2009; cf. Grodner et al. 2010, Breheny, Ferguson & Katsos 2013).

Thus, in view of the processing literature, an implicature approach to cleft exhaustivity is anything but straighforward. As Destruel & De Veaugh-Geiss (2018: p. 12) [Chapter 2: p. 86] write: "To the extent that we can take response time measures to represent indexes of processing difficulty, the finding that, in English, participants take less time to provide 'false' judgments in +VIOLATION visual contexts appears to be at odds with [. . .] much of the experimental literature reporting a delay in implicature computation". Thus, the authors conclude the exhaustive inference in clefts must not be pragmatic, at least not in a way comparable to scalar implicatures. However, a presuppositional analysis similarly poses problems. For instance, although a presuppositional analysis might be able to account for the cross-linguistic differences found between English and French, presuppositions should not be cancellable in unembedded environments, contrary to their results.

Further mysteries emerge when considering individual participants' responses for German (Chapter 1) and Akan (Chapter 4) (and alluded to for English in Footnote 11 in Destruel & De Veaugh-Geiss 2018 in Chapter 2). Specifically, a post-hoc investigation of the by-participant truth-value judgments in Chapter 1 (De Veaugh-Geiss et al. 2018) and in Chapter 4 (De Veaugh-Geiss in preparation) reveals striking parallels between clefts and definite pseudoclefts. What the authors report is that the mid-range judgments were driven by two different participant groups: roughly half of the population treated clefts and definite pseudoclefts exhaustively, choosing

to judge the sentences as 'false' when exhaustivity was violated, and the other half treated them non-exhaustively, choosing to judge the sentences as 'true' despite the violation. The authors conclude that the *semantic definite* approaches to cleft exhaustivity—that is, the theories which predict parallels between clefts and definite pseudoclefts (e.g., Percus 1997, Büring & Križ 2013, etc.)—were on the right track, but not quite right. They contend that exhaustivity must not be conventionally-coded as a presupposition (Percus 1997, Büring & Križ 2013) in the cleft/definite pseudocleft structure—as the semantic definite accounts predict—since then one would expect the exhaustive inference to arise robustly and systematically across participants and trials, which the data does not bear out.

Specifically regarding the homogeneity account of Büring & Križ 2013, Križ 2016, 2017—one of the theories in the aforementioned semantic definite accounts—Renans & De Veaugh-Geiss (2019) demonstrate in Chapter 3 that this approach to modelling cleft exhaustivity makes the wrong predictions when considering predicate interpretation. The *homogeneity approach*, in particular in Križ 2017, posits that a negated cleft as in '*It wasn't Carlos who biked.*' (as well as its positive counterpart) will be neither true nor false when the cleft pivot is properly contained in the cleft referent, e.g., Carlos and someone else biked—that is, when there is an exhaustivity violation. Homogeneity is the term for this truth-value gap, and crucially, it is not predicted to be sensitive to the distributive/non-distributive interpretation of the predicate. By contrast, the accounts lumped together as the *alternative-based approach* (among many others, Velleman et al. 2012, Renans 2016b; for details, see Križ 2017: §2 and Renans & De Veaugh-Geiss 2019: §2.3 [Chapter 3: §3.2.3]) predict that the acceptability of such sentences will in fact depend on the distributive vs. non-distributive interpretation of the predicate.

Renans & De Veaugh-Geiss (2019) exploit this divergence in predictions to test experimentally—in an acceptability-judgment task using a 1–7 ordinal scale—two broad camps in the cleft exhaustivity debate. For the sake of concreteness, an example trial from their study is found in (32) (here showing the non-distributive context only).

(32) a. *Non-distributive context*:
Carlos and Andrea biked together on a tandem bike.
b. *Target*:
It wasn't Carlos who biked. Carlos and Andrea biked.

The key idea here is that the predictions of the acceptability of (32b) in a non-distributive context such as (32a) depend on the theory. In a nutshell, the homogeneity approach predicts a truth-value gap since homogeneity is not sensitive to predicate interpretation, and thus the series of sentences in (32b) should be invariably

unacceptable; the alternative-based approach, on the other hand, predicts (32b) to be acceptable. What the authors found is that participants in fact judged the sentences in non-distributive contexts as acceptable contrary the predictions of the homogeneity approach.[18] Similar to the conclusions in De Veaugh-Geiss et al. 2018 (Chapter 1) and De Veaugh-Geiss in preparation (Chapter 4), albeit for wholly different reasons, Renans & De Veaugh-Geiss conclude that the homogeneity approach is not compatible with their results.

Although Renans & De Veaugh-Geiss (2019) (Chapter 3) do not suggest any particular approach to modelling cleft exhaustivity, De Veaugh-Geiss et al. (2018) (Chapter 1), Destruel & De Veaugh-Geiss (2018) (Chapter 2), and De Veaugh-Geiss (in preparation) (Chapter 4), following the dynamic account in Pollard & Yasavul (2016), propose that cleft (and pseudocleft) exhaustivity be modelled as a discourse-pragmatic phenomenon—and not an implicature as in previous pragmatic approaches (Horn 1981, 2014, De Veaugh-Geiss et al. 2015). I will leave the details for later chapters, in particular Section 1.4 in Chapter 1, but in brief, under this approach the exhaustive inference in clefts and pseudoclefts comes about via the anaphoric existence presupposition encoded in both sentence types. The basic idea is that by resolving the existence presupposition to the maximal discourse referent introduced by an (implicit) *wh*-question, an exhaustive interpretation arises; by contrast, if the existential is resolved to a non-maximal discourse referent such as that introduced by an indefinite antecedent, a non-exhaustive interpretation is obtained. De Veaugh-Geiss et al. (2018) (Chapter 1) and De Veaugh-Geiss (in preparation) (Chapter 4) argue that this approach can account for the inter-speaker variability found in the data: either a strongly exhaustive or a non-exhaustive interpretation may arise depending on the resolution strategy used by a hearer.

Destruel & De Veaugh-Geiss (2018) claim that this discourse-pragmatic approach can account for the cross-linguistic variability they report as well (see Section 2.5 in Chapter 2 for details). French differs from the other languages, in particular English, in the discourse-semantics of clefts. French clefts have a less stringent question-answer congruence requirement compared to English clefts (cf. Abrusán 2016 for question-answer congruence in clefts): that is, *c'est*-clefts can answer a range of QUDs, including all-new focus questions such as *What happened?* (Clech-Darbon, Rebuschi & Rialland 1999, Katz 2000, Lambrecht 2001). Since French

18 Note that this is a simplified picture. The study was interested in comparing target sentences as in (32b) in contexts with both a *non-distributive* interpretation, as in (32a), as well as a *distributive* interpretation, not shown here; for a detailed description of the study, see Chapter 3. Moreover, in the first version of the experiment, sentences such as (32b) were merely judged as comparatively *more* acceptable in non-distributive contexts than in distributive contexts; only in the second and third versions of the experiment—when the targets were presented as auditory stimuli—were the sentences in non-distributive contexts judged as acceptable.

clefts can be used to answer a broader range of questions, there are multiple avenues for resolving the anaphoric existence presupposition, with only one resulting in an exhaustive interpretation—thus, overall a less exhaustive interpretation is found. Finally, Destruel & De Veaugh-Geiss (2018) argue that for English the question corresponding to an exhaustive interpretation is the one most easily derivable from the focus-background structure of the cleft, and thus absent any context the exhaustive interpretation will be the initial, default interpretation[19]—accounting for the quicker response times for responses corresponding to exhaustive interpretations and slower response times for those corresponding to non-exhaustive ones.

The four papers presented here on German, English, French, and Akan are a modest and—I hope—engaging contribution to the debate on cleft exhaustivity. Without further ado, I turn now to the papers.

19 See also Titov 2019: §3.2.2 for claims that Akan *nà*-clefts 'favor' an exhaustive interpretation absent a context licensing otherwise.

Chapter 1

That's not quite it: An experimental investigation of (non-)exhaustivity in clefts*

Abstract We present a novel empirical study on German directly comparing the exhaustivity inference in *es*-clefts to exhaustivity inferences in definite pseudoclefts, exclusives, and plain intonational focus constructions. We employ mouse-driven verification/falsification tasks in an incremental information-retrieval paradigm across two experiments in order to assess the strength of exhaustivity in the four sentence types. The results are compatible with a parallel analysis of clefts and definite pseudoclefts, in line with previous claims in the literature (Percus 1997, Büring & Križ 2013). In striking contrast with such proposals, in which the exhaustivity inference is conventionally coded in the cleft-structure in terms of maximality/homogeneity, our study found that the exhaustivity inference is not systematic or robust in *es*-clefts nor in definite pseudoclefts: Whereas some speakers treat both constructions as exhaustive, others treat both constructions as non-exhaustive. In order to account for this unexpected

* We wish to thank Anna-Christina Boell and Julia Hensel for preparation of the German stimuli, and Friederike Buch, Ferdinand Kreuzkamp, Mareike Philipp, and Rico Winkel for their assistance in running the experiments at the Universität Göttingen and Universität Potsdam. This work has greatly benefited by the insightful comments and constructive criticisms of the editor and three anonymous reviewers at *Semantics & Pragmatics*, David Beaver, Judith Degen, Elsi Kaiser, and Maribel Romero as well as from the participants at *Experimental Pragmatics 2015* in Chicago, USA, the *XPrag.de* annual meeting in Göttingen, Germany, the *Experimental Semantics* workshop at ESSLLI 2015 in Barcelona, Spain, the *Experimental Pragmatics* workshop at the Societas Linguistica Europaea 2015 in Leiden, Holland, as well as the PhD reading group at the Universität Konstanz, among many others too numerous to list here. This research is funded by the German Research Foundation (DFG) as part of the project "Exhaustivity in *it*-Clefts" in the priority program XPrag.de. All errors are our own.

finding, we argue that the exhaustivity inference in both clefts and definite pseudoclefts—specifically those with the compound definite *derjenige*—is pragmatically derived from the anaphoric existence presupposition that is common to both constructions.

DOI: https://dx.doi.org/10.3765/sp.11.3

Keywords: experimental study; exhaustivity; *es*-clefts; definite pseudoclefts; anaphoric existence presupposition

1.1 Introduction

The sentence in (1) is the German counterpart of the English *it*-cleft provided in the translation. In this paper, we will simply refer to such German sentences as *clefts*, although they are only one of several possible cleft-structures in German and they are mainly known in the literature as *es*-Spaltsätze (*es*-Clefts) (Huber 2002, Altmann 2009). To define the terminology used here, we characterize clefts as constituted by the neuter pronoun *es*, which is arguably an expletive (Gazdar et al. 1985, Pollard & Sag 1994, É. Kiss 1999), a copula verb and a cleft pivot which agree in number and person, and a cleft relative clause with a relative pronoun which agrees in number and gender with the cleft pivot.[1]

(1) Es ist JOHN, der getanzt hat.
it COP.SG John REL.MASC.SG danced has
'*It is John who danced.*'

In addition to their so-called canonical inference, cleft sentences are frequently claimed to come with two inferences of particular interest for semantic theory: an existential inference and an exhaustivity inference. For example (1), we exemplify these inferences in (2).

(2) a. Canonical inference: John danced.
b. Existential inference: Somebody danced.
c. Exhaustivity inference: Nobody other than John danced.

1 In this study, we only consider focus-background clefts, in which the cleft pivot carries the nuclear focal pitch accent and the cleft relative clause is de-accented. As opposed to this, so called topic-comment clefts have also been observed in the literature (for a recent discussion, see Hedberg 2013); however, we have nothing to say about them in this paper.

There is little controversy in the literature about the canonical and the existential inferences in clefts. While the existential inference is typically assigned the status of a presupposition (e.g., Horn 1981, Rooth 1996, Delin 1992, Hedberg 2000) and commonly considered obligatory (Dryer 1996, Rooth 1999; cf. Büring & Križ 2013, discussed further in Section 1.4.1), the canonical inference is generally taken to be part of the proffered content; that is, it is an at-issue semantic inference. By contrast, the interpretive status and obligatory presence of the exhaustivity inference is very much debated in the literature.[2]

The semantic literature offers two main sources for the origin of the exhaustivity inference. On the one hand, there is a pragmatic account of clefts, which was first proposed in Horn 1981 and more recently defended in Horn 2014. According to the pragmatic account, cleft exhaustivity is not conventionally coded in the cleft structure itself but is rather a generalized conversational implicature, derived from the fact that clefts also have an existential presupposition. There are variants of this idea which mainly build on the observation that the cleft pivot is focused. For instance, De Veaugh-Geiss et al. 2015 argue that the source of pragmatic exhaustification in clefts lies in the non-canonical, unambiguous focus marking. A similar conclusion is put forward by the dynamic account of Pollard & Yasavul 2016, according to which *it*-clefts in English are anaphoric expressions that specify their anaphoric antecedent, whereby exhaustivity occurs as part of a question-answer paradigm.

At the same time, many scholars assume a semantic source of cleft exhaustivity, in which the exhaustive inference is conventionally coded in the cleft structure itself. A large part of such semantic accounts builds on the connection between definiteness and clefts and is thus dubbed by us the *semantic definite* account of clefts; see Akmajian 1970, Szabolcsi 1994, Percus 1997, Büring & Križ 2013. Such accounts hold that cleft sentences such as (1) contain a covert determiner element, or some more complex compositional derivation, that makes them semantically equivalent to definite descriptions such as (3), which in turn are assumed to be semantically exhaustive. In particular, the exhaustivity inference is typically modeled as a maximality presupposition (Percus 1997) or as a homogeneity presupposition (Büring & Križ 2013).

(3) Derjenige, der getanzt hat, ist JOHN.
DEF.MASC.SG REL.MASC.SG danced has COP.SG John
'*The one that danced is John.*'

2 The title of this article, "That's not quite it: An experimental investigation of (non-) exhaustivity in clefts" is a reference to Büring & Križ's (2013) *Semantics and Pragmatics* article "It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites)," which we take issue with here, in particular the claim that *it*-clefts are semantically exhaustive.

A further type of semantic account has recently emerged which suggests that exhaustivity in English clefts, and arguably in German clefts as well, is derived from a conventional interaction between clefts and the question under discussion (sensu Roberts 2012), hence being closer to the focus-based pragmatic accounts. Such accounts include Velleman et al. 2012, Destruel et al. 2015, Beaver & Onea 2015.

Even though this is common practice in the literature, we will not refer to examples like (3) simply as definite descriptions. Example (3) is a specificational construction involving a heavy definite description consisting of a complex definite determiner *der*.MASC / *die*.FEM / *das*.NEUT *–jenige* (involving the distal demonstrative stem *jen*) and a full relative clause. We will call such constructions *definite pseudoclefts*. We choose this theory-neutral label in order to signal two facts. First, the structure in (3) involves a clear definite description on the surface, which certainly plays a role in the semantic interpretation of such structures. But, second, the structure is related but not identical to German pseudoclefts, such as the one in (4), which are also known as *wh*-clefts. In German, such pseudoclefts are built around a *wh*-element in a free relative clause. Crosslinguistic observations, however, show that there are languages in which pseudoclefts obligatorily appear with a definite article, such as Romanian or Spanish (Romero 2005).

(4) Wer am besten getanzt hat, ist JOHN.
who the best danced has COP.SG John
'*Who danced the best is John.*'

All three inferences shown in (2) for clefts are also typically attributed to definite descriptions in general, and to definite pseudoclefts like the one in (3) in particular. As with clefts, these inferences have been hotly debated for definite descriptions too. In particular, the presuppositional status of both the uniqueness inference (e.g., Szabo 2000, Ludlow & Segal 2004) and the existential inference (Coppock & Beaver 2015) has been challenged. Still, the mainstream view seems to be that definite descriptions presuppose both existence and uniqueness.

We do not have much to add in this paper to the debate about the existence and uniqueness presuppositions of definite descriptions. What we find is that the literature has never considered that definite pseudoclefts differ from more run-of-the-mill definite descriptions in any of these respects. Instead, scholars pointed out parallels between clefts and definite pseudoclefts, also as far as exhaustivity is concerned. They assume this to indicate that clefts are like definites in general. A particularly striking example for this is Križ 2015. Anticipating the discussion to come, we will question this assumption in the discussion of our experimental findings and suggest that clefts may well be similar to definite pseudoclefts without apparently sharing properties of definiteness. In a nutshell, we claim that the exhaustivity inference in definite pseudoclefts occurs independently of the uniqueness presupposition of

the definite article, namely as a by-product of resolving an anaphoric existence presupposition.

One general contention in the theoretical literature is whether cleft exhaustivity is conventionally coded and thus clefts should invariably give rise to exhaustivity inferences. This contention has recently been challenged by findings of several experiments on the interpretation of clefts. Studies such as Onea & Beaver 2009, Destruel et al. 2015, De Veaugh-Geiss et al. 2015, Destruel 2012, Byram-Washburn, Kaiser & Zubizarreta 2013 and others report various ways in which cleft exhaustivity does not align with standard expectations raised by semantic inferences of any kind. In other words, the experimental findings point to a pragmatic analysis of the exhaustivity inference in which exhaustivity is not conventionally coded as part of the literal semantic meaning of clefts.

The main objective of this article is to bridge the empirical gap opening up between the vast majority of theoretical accounts, on the one hand, and the experimental findings, on the other, by presenting and discussing the results of two novel experimental studies on the nature of cleft exhaustivity. The experiments come in the form of verification and falsification tasks in an incremental information-retrieval paradigm, akin to the *incremental verification task* (IVT) used by Conroy (2008) to test for the available interpretations of scopally ambiguous strings and by Franke, Schlotterbeck & Augurzky (2016) for literal, local, and global readings of embedded scalars. Our studies improve upon existing experimental studies on *it*-clefts in systematically comparing the interpretation of clefts as in (1), definite pseudoclefts as in (3), plain intonational focus constructions as in (5), and exclusives as in (6).

(5) JOHN hat getanzt.
John has danced.
'*John danced.*'

(6) Nur JOHN hat getanzt.
Only John has danced.
'*Only John danced.*'

This four-way comparison including uncontroversial instances of pragmatic exhaustivity (plain accent focus) and truth-conditional exhaustivity (exclusives) leads to a more complete view of the problem, and it allows to test for the predictions of a large array of different theories.

Our experimental results show, somewhat surprisingly, that the pragmatic implicature account as well as the semantic definite account are both right *and* wrong to a certain extent. As predicted by the definite account, clefts were interpreted exactly like definite pseudoclefts in the experiments, contrasting with plain foci and exclusives. Conversely, unlike what is predicted by the definite account and other semantic analyses of exhaustivity inferences in clefts, neither clefts nor definite pseudoclefts are obligatorily interpreted as exhaustive. This finding seems to call for a pragmatic account that treats the exhaustivity of clefts and of definite pseudoclefts in parallel ways.

In our analysis of the experimental findings, we suggest a version of Horn's (1981) original analysis, in which the exhaustivity of clefts rests essentially on their existential presupposition. While Horn (1981) had to stipulate a specific pragmatic rule in order to derive exhaustiveness from the existential presupposition, our proposal derives exhaustivity from the anaphoricity of clefts: Exhaustivity arises whenever the anaphoric antecedent of the existential presupposition is interpreted as maximal by the hearer, in a way similar to what Pollard & Yasavul (2016) propose. Crucially, our pragmatic account deviates from the earlier pragmatic implicature account in De Veaugh-Geiss et al. 2015 in that it does not derive an exhaustivity implicature from the explicit and unambiguous structural marking of focal alternatives (cf. Büring [2015]). Moreover, we propose a similar analysis for definite pseudoclefts with *derjenige*, suggesting they differ from plain definite descriptions in that they must be interpreted as obligatorily anaphoric expressions and, moreover, these constructions pragmatically derive exhaustivity independently of the maximality semantics of the compound definite article.

The paper is structured as follows: In Section 1.2 we discuss previous theoretical accounts and their predictions for the behavior of clefts and definite pseudoclefts in semantic experiments. We also provide an overview of previous experimental studies on cleft exhaustivity, together with a brief discussion of their shortcomings. We then introduce our own experimental approach and show why it has advantages over previous approaches. Section 1.3 and Section 1.4 form the heart of the paper. In Section 1.3 we describe the experimental set-up and results in detail. In Section 1.4 we put forward a pragmatic analysis in terms of the anaphoric existence presupposition of clefts and of definite pseudoclefts with *derjenige*, and we will demonstrate how this analysis can account for the experimental data. In Section 1.5 we conclude with a summary of the main findings and a brief discussion.

1.2 Theoretical and experimental approaches

In this section we will briefly introduce the predictions of the main theoretical approaches to cleft exhaustivity. We will also sum up the existing experimental findings and discuss some shortcomings of previous experimental studies. These shortcomings will motivate the experimental studies to be presented in the next section. In our studies, we focus on two important interpretive aspects, which constitute the central parameters against which we will evaluate experimental and theoretical approaches: (i) The *strength* of the exhaustivity inference across experimental conditions and speakers; and (ii) the *parallel* behavior of clefts and definite pseudoclefts regarding exhaustivity. The term *strength* is a cover term that refers to the overall robustness and systematicity of an exhaustivity inference. A robust inference is both obligatory

and non-cancellable across all contexts for all speakers of a language group. The term systematicity is related to the notion of robustness but it refers specifically to the regularity of exhaustivity across experimental set-ups, experimental conditions, and also across speakers.[3] Regarding the question whether exhaustivity in clefts is semantic or whether it is pragmatic, strength seems to be a key feature.

The second parameter, *parallel* behavior, is included because of the important research tradition that derives cleft exhaustivity from an underlying definite structure, which has been assumed to be intimately related if not identical to the structure of definite pseudoclefts. If such approaches, including the recent proposal in Büring & Križ 2013, are on the right track, clefts and definite pseudoclefts are expected to behave in fully parallel ways regarding exhaustivity. Importantly, the parallelism parameter first and foremost touches upon the question of how cleft (and pseudocleft) exhaustivity is structurally derived, and only indirectly upon the question of whether the exhaustivity inference is semantic or pragmatic in nature. Given the widespread contention in the literature that definites are semantically exhaustive, a parallel behavior of clefts and definite pseudoclefts could indeed be taken as evidence for semantic exhaustivity in clefts. However, our experimental findings will suggest that this hypothesis cannot be maintained, and, consequently, that the exhaustivity inference is not semantic in clefts, nor is it in definite pseudoclefts.

In the following, we will first consider the predictions of three types of theories regarding the two dimensions of possible variation. Building on these predictions, we then present some insights from existing experimental data.

1.2.1 Theoretical predictions

Theoretical approaches to cleft exhaustivity divide into the non-conventionally-coded pragmatic and the conventionally-coded semantic accounts. The most prominent pragmatic account is the implicature analysis in Horn 1981, 2014. According to Horn, the implicature is triggered by the interaction of the obligatory existence presupposition of clefts and the additional use of a non-canonical and less economical cleft structure. In particular, Horn 1981 proposes an idiosyncratic, structure-specific pragmatic principle of derivation according to which if a speaker uses an *it*-cleft of the form *it is* α *that* P which presupposes $\exists x.P(x)$ and asserts $P(\alpha)$, then she implicates $\forall x.x \neq \alpha \rightarrow \neg P(x)$ in the form of a generalized conversational implicature.

3 Note that robustness and systematicity may occur independently of each other: A systematic inference can be context dependent but still uniform across speakers; however, in this case, it would not be a robust inference. In other words, a robust inference occurs in any context and it is non-cancellable, though possibly only for a sub-group of the population; a systematic inference can be context-dependent, detected with various experimental methods and, most importantly, it is not only attested with a (small) part of the speakers of the respective language.

Regarding the two parameters ±*strength* and ±*parallel* discussed above, the pragmatic implicature account predicts the exhaustivity inference to be subject to cancellability or variability across contextual conditions and speakers. Horn (1981) provides naturally occurring examples such as (7) as points in case. However, Horn (1981: 133, ex. 18d) also notes that out-of-the-blue cancellation is not always possible, as shown in (8). The explanation given by Horn (1981) for cases in which cancellation of the implicature appears to be difficult or impossible is that in such cases the fact that the speaker uses a marked structure (the cleft as compared to the canonical sentence) would not be justified if exhaustivity did not hold.

(7) It's the ideas that count, not just the way we write them.
[Richard Smaby, lecture, *via* Ellen Prince; Horn 1981 (13d)]

(8) #It was a pizza that Mary ate; indeed it was a pizza and a calzone.

So, the main prediction of the pragmatic account is a lack of strength in the exhaustivity inference. There is no clear prediction regarding the parallel behavior of definite pseudoclefts, for the main reason that Horn's theory of cleft exhaustivity makes no claim about definite pseudoclefts at all. However, Horn & Abbott (2016) clearly argue that uniqueness is a conventional part of the meaning of definite descriptions; that is, in their terminology, it is a conventional implicature. This predicts a robust and systematic exhaustivity inference to obtain with definite descriptions, and hence no parallel behavior between clefts and definite pseudoclefts. While we consider this a prediction of Horn's theory, we note in passing that a possible parallel behavior between clefts and definite pseudoclefts in itself does not show that Horn's pragmatic theory of cleft exhaustivity is misguided. It could also be taken to show that his analysis is incomplete, and that Horn & Abbott's analysis of definite descriptions, is, independently, incorrect or at least it does not apply to definite pseudoclefts. More importantly, however, if clefts and definite pseudoclefts behave in a parallel fashion, Horn's theory must be extended to explain this fact irrespective of Horn & Abbott's views on definiteness in general.

Next to the pragmatic analysis, there are two prominent semantic approaches to cleft exhaustivity. The semantic definite accounts treat clefts and definite descriptions as sharing the logical form of identity statements in which a discourse referent is identified with the cleft pivot or restrictor predicate (e.g., Percus 1997, Büring & Križ 2013, Križ 2017). More specifically, in Percus 1997 clefts contain a covert definite operator and have the underlying syntax and semantics of a definite description, whereas in Büring & Križ 2013 clefts and definite descriptions can be treated in parallel in terms of their semantic contribution, although the analysis for clefts does not strictly depend on this. The exhaustivity inference in these approaches is either modeled in terms of a maximality presupposition (Percus 1997, Szabolcsi

1994)[4] or a homogeneity presupposition (Büring & Križ 2013). For an example like (1), the maximality account presupposes a maximal discourse referent that dances and asserts that this referent is identified with the pivot. By contrast, assuming a homogeneity presupposition, it is asserted that John dances and it is presupposed that John is not a proper mereological subpart of the maximal individual that danced; i.e., either nobody danced or John is the maximal individual that danced.[5] Most clearly for Percus 1997, and potentially for Büring & Križ 2013 as well, definite pseudoclefts as in (3) are expected to share with their cleft counterparts the asserted and presupposed meaning. Regarding the two main parameters discussed above, the predictions of the definite semantic account for cleft exhaustivity are the clearest of all. If clefts may be considered definites in essence and if both are assumed to conventionally-code exhaustivity, clefts and definite pseudoclefts are expected to show parallel interpretive behavior.[6] In particular, both sentence types are predicted to exhibit exhaustivity inferences in a robust and systematic manner.

Note, however, that definite descriptions are not a homogeneous class as far as exhaustivity is concerned. Some definites seem to be less exhaustive than others; e.g., weak definites (Schwarz 2009, Barker 2004, Carlson et al. 2006) or seemingly indefinite definites (Carlson & Sussman 2005) do not presuppose uniqueness. In any case, it is not obvious whether it is possible to treat all different kinds of definites alike. Other approaches such as Abbott (2014) distinguish semantic uniqueness and referential uniqueness, defining the latter as follows: "[T]he essence of definiteness in a definite description is that the speaker intends to use it to refer to some particular entity, and (crucially) expects the addressee to be able to identify that very intended referent." This pragmatic notion of referential uniqueness incorporates the idea that uniqueness may refer to the discourse status of previously mentioned discourse referents or discourse referents entailed by the preceding discourse. It allows for the use of definite descriptions with familiar (Heim 1982) rather than semantically unique referents, as long as they are identifiable in discourse. The semantic approach to clefts described above, however, analyzes definites, and in particular definite

4 Although Percus (1997: 342) refers to a uniqueness presupposition, we think it is more appropriate to refer to maximality since Percus's analysis would apply to both singular and plural entities. In particular, Percus explicitly entertains the possibility that the covert NP-proform involved in clefts is unspecified for number, thereby allowing for both singular and plural reference.

5 For details on the theoretically-predicted presupposition failure when a plurality such as John and Mary danced, see Büring & Križ 2013 and Križ 2017.

6 However, since in Büring & Križ 2013 the analysis of cleft exhaustivity in terms of homogeneity does not depend on the same analysis being extended to definite descriptions, exhaustivity violations as in our experiments may not predict full parallelism between the two sentence types: the cleft will incur a homogeneity violation (assuming a homogeneity presupposition), the singular definite a uniqueness violation (assuming a uniqueness presupposition). We thank an anonymous reviewer for pointing this out.

pseudoclefts, as presupposing uniqueness. In section 1.4.2, by contrast, we will argue that definite pseudoclefts do not fall into the category of semantically unique definites and adopt a familiarity analysis instead.

The second prominent semantic account of cleft exhaustivity is the inquiry-terminating (IT) construction analysis of Velleman et al. (2012), in which clefts have a semantically predicative form just as their canonical counterparts, with an additional meaning component giving rise to exhaustivity. In this analysis, cleft structures are treated as conventional devices to give a final and therefore complete answer to a question. In particular, they factor the meaning components of clefts into two components of different discourse-semantic status. At the at-issue level (e.g., Simons et al. 2010, Tonhauser et al. 2013), a cleft asserts the same as the respective canonical sentence would, namely that the predicate denoted by the cleft relative clause holds of the cleft pivot. At the same time, clefts express the not-at-issue inference that all stronger focus alternatives to the cleft prejacent are excluded. The at-issue truth of the prejacent and the exclusion of stronger alternatives are modeled by means of MIN- and MAX-operators, as shown in (9) for the cleft in (1). In this account, clefts have the same semantics as sentences with exclusives (*only*) except for the important difference that with exclusives the at-issue and not-at-issue status of the two components is reversed.

(9) It is John who danced.

At-Issue: MIN(⟦JOHN danced⟧) = There is a focus alternative that is at least as strong as the proposition *John danced* which is true.

Not-At-Issue: MAX(⟦JOHN danced⟧) = All stronger focus alternatives entailing *John danced* (e.g., *John and Bill danced*; *John, Bill, and Mary danced*; etc.) are false.

This account makes a clear prediction about the strength of exhaustivity, which is similar to the prediction of the definite account above: Exhaustivity in clefts is expected to be systematic and robust. We do not know of any case where the meaning of definites is treated by means of focus-sensitive MIN- and MAX-operators, and thus, the IT-construction account does not make any predictions about the parallel behavior of clefts and definite pseudoclefts with regard to exhaustivity.

Summing up, the (A) pragmatic, (B) semantic definite, and (C) semantic IT-construction approaches differ in their predictions regarding the two parameters identified above, i.e., [± strength] (robustness and systematicity) and [± parallel] with respect to definite pseudoclefts. The predictions of each are schematically presented in Table 4.1. Strikingly, Table 4.1 contains only three of the four possible combinations of values for the two parameters. There is one logically possible combination that is not predicted by any formal account of cleft exhaustivity: [–

strength] and [+ parallel]. On this setting, cleft sentences are expected to behave like definite pseudoclefts, but, crucially, the interpretive effect would not be robust nor systematic, but rather subject to contextual factors, experimental conditions, or inter-speaker variability. Eventually, we report in Section 1.3 that it is this hitherto unpredicted combination of parameter values which we find in our experiments. In other words, our experimental findings will show all existing formal theories of exhaustivity in clefts and definite pseudoclefts to be wrong, at least in part. We first turn to existing experimental research on the topic.

	± *strength*	± *parallel def.pse.*
(A) *pragmatic*	−	−
(B) *semantic definite*	+	+
(C) *semantic IT-construction*	+	+/−

Table 1.1 *Predictions of three theoretical approaches to cleft exhaustivity.*

1.2.2 Existing experimental approaches

Recent years have seen an increase in experimental approaches to the interpretation of cleft sentences in English, German, and French. For the most part, the experimental studies were motivated by the fact that the theoretical literature was incapable of settling the exact interpretive status of the exhaustivity inference on the basis of pure introspection and native speaker intuitions. One problem is that intuitions on cleft exhaustivity are often too shaky and variable, necessitating the need for controlled and quantifiable experimental methods; another problem is that different theories tend to focus on different subsets of the data and to disregard others, necessitating a more comprehensive study on the relevant aspects of exhaustivity. Notably, while the formal linguistic literature exhibits a preference for semantic analyses of cleft exhaustivity, all existing experimental studies point toward a pragmatic analysis, in line with the pragmatic implicature analysis of Horn 1981, 2014.

The study of Onea & Beaver 2009 (and replications thereof) used the *Yes, but...*-test comparing clefts, exclusives, and canonical sentences. They found that participants chose weaker continuations whenever exhaustivity was violated in a cleft, as compared to exclusives in which exhaustivity is at-issue. These findings indicate that the exhaustiveness of clefts is weaker than would be expected on a semantic account. However, Mayol & Castroviejo (2013) and Xue & Onea (2011) claim that corrective *but*-responses are in fact contradictions of not-at-issue content in the sense of Simons et al. 2010 and Tonhauser et al. 2013. Hence, the results of

Onea & Beaver 2009 just show that exhaustivity in clefts is not-at-issue, but would be in line with a pragmatic as well as a semantic account.

Byram-Washburn, Kaiser & Zubizarreta (2013) used written material and a dialogue-setting for testing the acceptability of clefts comparing exhaustivity violations and violations of contrastiveness inferences, which are also often attributed to clefts (Destruel & Velleman 2014, Destruel, Beaver & Coppock 2017). They found that a violation of contrastiveness leads to much lower acceptability ratings than the violation of exhaustiveness. Hence, they argue that cleft exhaustivity is not a semantically-coded presupposition, but rather a conversational implicature. They are, however, missing a direct comparison with maximality presuppositions, while the presumed interpretive status of the contrastiveness inference as a presupposition is not independently assessed.

Finally, De Veaugh-Geiss et al. (2015) report the results of an acceptability study. The aim of this study was to clarify whether or not the difference in at-issueness between the canonical inference and the exhaustivity inference of clefts is sufficient to explain the apparent weakness of the exhaustivity inference observed for clefts. The study showed that the exhaustivity inference in clefts is easier to cancel than, for instance, the prejacent of exclusive particles (*only*), even though both meaning components are commonly treated as not-at-issue (Horn 2014). Note that the acceptability ratings for exhaustivity cancellations in clefts were still quite low, though, with judgments in the mid-range of a 7-point scale. In a follow-up experiment with definite pseudoclefts in place of *it*-clefts, cancellations of uniqueness were by contrast treated in the same way as cancellations of the prejacent of exclusives, and again, both inferences are not-at-issue. Hence, De Veaugh-Geiss et al. (2015) argued that at-issueness cannot be the sole factor responsible for the observed weakness. Rather, the experimental findings were taken in support of a pragmatic implicature account of cleft exhaustivity.

Summing up, the previous experimental studies have delivered ample evidence for the different status of the exhaustivity inference in clefts, on the one hand, and the at-issue exhaustivity expressed by exclusive particles, on the other. The experiments also provide some evidence in favor of a pragmatic nature of the exhaustivity inference, which mostly comes in the form of weakening effects (cancellability) and its sensitivity to contextual factors (non-robustness). At the same time, however, the experimental results do not provide conclusive evidence for the pragmatic implicature analysis of cleft exhaustivity. Either there are problems with linking the experimental findings to the nature of pragmatic or semantic inferences (Onea & Beaver 2009); or the exhaustivity effect in clefts is compared to inferences of equally unclear semantic or pragmatic status (Byram-Washburn, Kaiser & Zubizarreta 2013); or the graded values on the acceptability judgment scale are not as high as might be expected on a pragmatic account in which exhaustivity should be defeasible (De Veaugh-Geiss

et al. 2015). Moreover, most experimental studies fail to make a direct comparison of the exhaustivity effect in clefts with the maximality presupposition of definites, even though the latter is considered the most likely semantic source of cleft exhaustivity, at least according to large parts of the theoretical literature. Given this state of affairs, we conclude that a more systematic experimental study directly aimed at examining the relevant parameters listed in Section 1.2.1 is required. In the next section, we describe the experimental setup of such a study.

1.3 The experiments: Verification and falsification in an incremental information retrieval setting

In light of the above discussion, we propose a novel experimental approach to the study of exhaustivity inferences in clefts. Our approach departs from its predecessors in two ways: First, our experiments involve an explicit comparison between clefts and definite pseudoclefts, as well as a comparison with two well-established control constructions, exclusives and plain intonational focus (i.e., focus marked *via* a pitch accent). The explicit comparison between clefts and definite pseudoclefts should provide useful evidence for establishing whether the source of the exhaustivity inference is the same in both structures or not.[7] The explicit comparison with the control structures, and in particular with the plain focus condition, should provide evidence for establishing whether the exhaustivity inference is pragmatic in nature or not. Second, in order to overcome the observed difficulties in the interpretation of gradient acceptability ratings, we use an incremental information retrieval paradigm that involves decision-making and interpretation procedures, namely verification and falsification with the option of continuation. Given the different kinds of tasks involved in verification and falsification of inferences, we also directly test for the strength of the inference.

Recently, variations on verification and falsification tasks have been employed by Abrusán & Szendrői 2013 in a truth-value judgment task on reference failure for definite NPs, and by Romoli & Schwarz 2015 in a covered-box paradigm on local accommodation of presuppositions (or, conversely, contexts with global presupposition failure) for the trigger *stop* when embedded under negation. We extend such experimental methods here to two classes of (alleged) definite expressions, namely definite pseudoclefts and cleft sentences. Although experimental studies with presupposition contradiction are few, the above studies have in fact found that

7 To avoid confusion, we should mention that when designing the experiment we shared the assumption made in the literature that definite pseudoclefts are plain definite descriptions and that they are, therefore, semantically exhaustive just like plain definites. Given the results of the experiments and further evidence we collected when analyzing the data, we consider this an inappropriate premise.

such sentences result in a majority of 'false' judgments (despite a 'can't say' option) (Abrusán & Szendrői 2013) and broad rejections (e.g., by selecting the covered box) (Romoli & Schwarz 2015). Moreover, verification and falsification experiments give rise to categorical judgments, which should in principle allow for an easier identification of non-gradient differences between the two target structures at hand. Finally, our experiments exhibit other important design features that allow for a controlled and systematic study of exhaustivity inferences in clefts and definite pseudoclefts.

i. The experiments explicitly control for at-issue semantic exhaustivity triggered by exclusive particles and for *bona fide* pragmatic exhaustivity triggered by instances of *in situ* prosodic focus in auditory stimuli.

ii. The experiments explicitly control for domain restriction in order to rule out any attempts at explaining exhaustivity violations away in terms of a subsequent enlargement of the quantificational domain.

iii. The experiments involve proper names referring to four individuals without additional specifications. Hence, there is no ordering of alternatives in terms of informational strength (for instance, scalar items such as *all* being logically stronger than *some*; see the extensive literature on scalar implicatures) in order to rule out attempts at explaining exhaustivity effects away by recourse to ordering on a logical or contextually-supplied scale.

1.3.1 Method

In this section we provide a general overview of the two mouse-guided sentence-picture verification/falsification experiments, which provide the empirical substance of this paper. Since the timeline and the stimuli in the two experiments were identical, we present the experiments together.

For Experiment I, we tested 32 native speakers of German, all students in Potsdam and Berlin, Germany (24 female, 8 male; average age: 25.6; age range: 20–48).[8] For Experiment II we tested 32 native speakers of German—distinct from Experiment I—mostly students in Göttingen, Germany (20 female, 12 male; average age: 27.8; age range: 19–52). The experiments took part in a laboratory environment using self-programmed Python scripts (PSFL: GNU/Linux v.3.4.2; Windows v.3.3.5)

8 There were 33 participants in Experiment I, but one participant was removed for erratic judgments on the exclusive control condition: 2 'continues,' 2 'false,' and 3 'true' judgments as well as one missing data point at Box 2 due to already having made a judgment at Box 1; cf. the highly consistent decision to continue in the exclusive condition in Experiment I for the remaining 32 participants, seen in the left graph in Figure 1.2 on page 48.

Auditory stimulus, e.g.: "It is Max who mixed a cocktail."

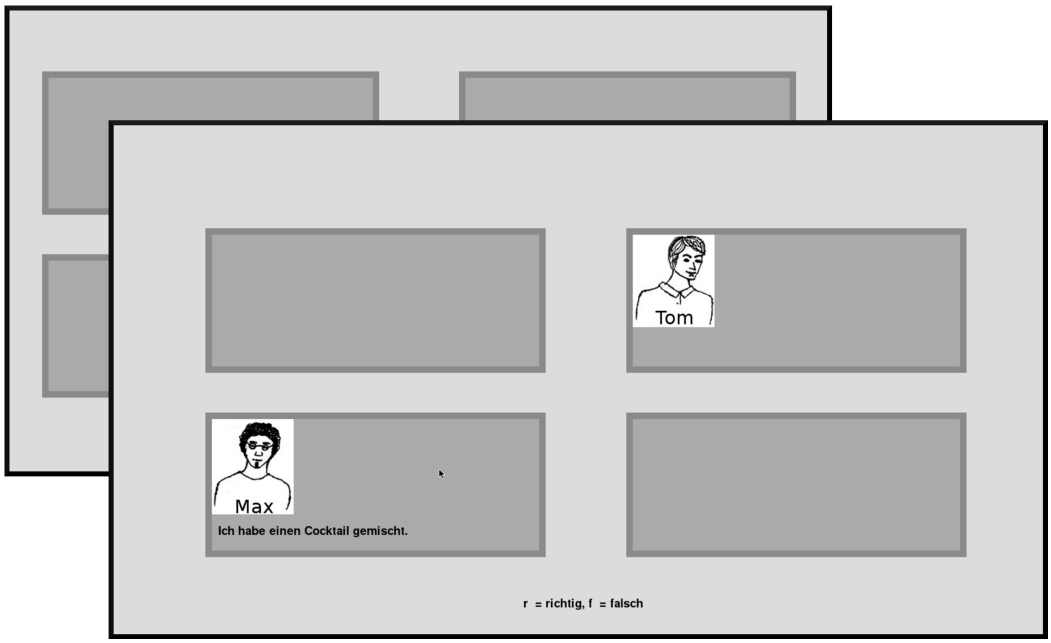


Figure 1.1 *(i) (behind) Start of each trial* *(ii) (front) Uncovering the 2nd box*

with the PyGame module (LGPL: v.1.9.2a0, Shinners 2011). Participants were compensated for their time.

In the instructions to the experiments, participants were introduced to four roommates: Tom, Max, Jens, and Ben. Participants were told that these roommates undertake various activities together. At the start of each trial, participants were presented with a computer screen showing four covered boxes while an audio stimulus played in their headphones. The screen appeared as in picture (i) in Figure 1.1. After hearing the stimulus, participants were asked to uncover as many boxes as necessary to decide if the audio sentence they heard was true or false. Each box contained an illustration of one of the roommates and a written first person statement about which action this roommate carried out, as in picture (ii) in Figure 1.1, in which Max says *Ich habe einen Cocktail gemischt* "I mixed a cocktail" in the bottom left box. At any time, participants could press *r* on the keyboard to signal that the sentence is *richtig* 'correct' or *f* to signal that the sentence is *falsch* 'false.' At Boxes 1–3 participants also had the choice of continuing by uncovering the next box.

Participants uncovered the boxes by moving the mouse over them. After entering the box, the cursor could not exit the box for at least 2000 ms. This procedure

was intended to keep participants from the unnecessary uncovering of too many boxes, such as, e.g., by automatically mousing over all four boxes and then making a judgment. When the cursor eventually left a given box, the text disappeared while the picture remained visible, although it was possible to move the mouse back into an uncovered box at any point of the trial to see the text again.[9] Hence, in picture (ii) in Figure 1.1, the participant had already uncovered the top right box which presented information that is no longer visible, and is currently viewing the bottom left box.

Although participants were free to choose which box they uncovered, they did not know that their choice had no influence on what they saw: The order of uncovering in the experimental setup for each trial was pre-determined, and which location they uncovered did not matter. This was done to prevent any strategies when it came to revealing contextual information. After participants made a judgment, the boxes onscreen were re-covered and the next target or filler item played in their headphones.

Stimuli and presentation Both experiments began with three practice trials to make sure that the participants understood how to control the mouse with respect to the contextual information onscreen, and that their task was to uncover just as many boxes as necessary. If participants uncovered too many boxes in the practice trial, they were reminded not to do so.

For both experiments, the auditory test items consisted of 32 target stimuli and 32 filler items, all in German. The target sentence varied in four sentence-type levels involving (i) clefts, (ii) definite pseudoclefts, (iii) exclusives, and (iv) plain intonational focus constructions, as shown in (10)–(13).

(10) Es ist MAX, der einen Cocktail gemischt hat.
it is Max who a cocktail mixed has
'*It is Max who mixed a cocktail.*' CLEFT

(11) Derjenige, der einen Cocktail gemischt hat, ist MAX.
the.one who a cocktail mixed has is Max
'*The one who mixed a cocktail is Max.*' DEF. PSE.

9 As pointed out by an anonymous reviewer, this makes the memory load heavier for participants, which may discourage continuing. Indeed, together with the 2-second delay, having the text disappear was intended to make unnecessary uncovering unfavorable. By having the text disappear, participants had to either commit this information to memory or re-uncover this information, thereby requiring an extra step backward; thus, a strategy of revealing unneeded information and only later making a judgment came at a higher cost. The control conditions ensure that, despite these costs, our measurements are nonetheless reliable.

(12) Nur MAX hat einen Cocktail gemischt.
only Max has a cocktail mixed
'*Only Max mixed a cocktail.*' EXCLUSIVE

(13) MAX hat einen Cocktail gemischt.
Max has a cocktail mixed
'*MAX mixed a cocktail.*' PLAIN FOCUS

In each of the four sentence types, the target sentence gives rise to an exhaustivity inference and a canonical inference, as discussed in the previous sections. These inferences are spelled out in (14).

(14) a. Exhaustivity: Nobody out of Tom, Jens, and Ben mixed a cocktail.
b. Canonical inference: Max mixed a cocktail.

There were 32 lexicalizations, with 8 per sentence type distributed in a Latin square design across 4 lists and randomized during presentation. For the targets, grammatical subjects were proper names and grammatical objects were non-specific indefinite determiner phrases with an unspecific interpretation that either referred to an inanimate object or an animal. The reason for using non-specific indefinite object determiner phrases was to avoid any confounding uniqueness effects from additional definite articles in the clause. The non-specific construal was ensured by the absence of narrow pitch accent on the indefinite determiner in the auditory stimulus.

In the definite pseudocleft sentences, the complex definite forms *derjenige*, *diejenige*, and *dasjenige* are compounds of the singular determiner elements *der-* 'the.MASC,' *die-* 'the.FEM,' or *das-* 'the.NEUT' plus *-jenige*, the latter derived etymologically from the demonstrative marker *jene/jener/jenes* meaning 'that one (over there).' For all stimuli in the definite pseudocleft condition, the complex definite in subject position was singular and masculine, and it displayed singular nominative marking and gender agreement with the masculine proper name in predicative position.

Given one of the auditory stimuli in (10)–(13), Table 1.2 gives example stimuli in English for each possibility crossing all the factors for Experiment I and Experiment II. The different factors will be presented in the following.

Factorial design of Experiment I Experiment I involved a 4*2 factorial design, the two factors being SENTENCE TYPE and EXHAUSTIVITY. The EXHAUSTIVITY factor has two levels: [+EXH] and [–EXH]. In the [+EXH] condition no box provides information that would violate the exhaustivity inference triggered by the target sentence. Hence, for our example, Tom, Jens, and Ben report having performed other actions than having mixed a cocktail. By contrast, in the [–EXH] condition the

	Exp. I (verifier)	**Exp. II (falsifier)**
Box 1	(*irrelevant information*) Jens: "I opened a bottle."	
Box 2 (EARLY RESPONSE)	(***canonical verified***) Max: "I mixed a cocktail."	(***exhaustivity falsified***) Ben: "I mixed a cocktail."
Box 3 / Box 4 (LATE RESPONSE)	**[+EXH]** (*exh. verified*) Tom/Ben: "I fetched a straw." or **[–EXH]** (*exh. falsified*) Tom/Ben: "I mixed a cocktail."	**[+CAN]** (*can. verified*) Max: "I mixed a cocktail." or **[–CAN]** (*can. falsified*) Max: "I fetched a straw."

Table 1.2 *Conditions of Experiment I (verifier) & Experiment II (falsifier).*

third or fourth box uncovered by the participant provides a piece of information that contradicts the exhaustivity inference. In this case, for instance, the third box would contain the picture of Jens reporting that he (also) mixed a cocktail (see Table 1.2).

Dependent variables of Experiment I In Experiment I, the second box which was uncovered always entailed that the canonical inference triggered by the target sentence was true; that is why it is called "verifier" in Table 1.2. Hence, for Experiment I as shown in Table 1.2, the second box explicitly reveals that Max mixed a cocktail, which is identical to the canonical inference of (10)–(13). The first box never contained any information that would be relevant to the canonical or exhaustivity inference of the target sentence (see Table 1.2).

With this background, we measured two dependent variables. The first dependent variable was the response immediately following the uncovering of the second box, which had three possible values, i.e., whether the participant judged the sentence true or false immediately or opted to continue by uncovering one or more further boxes. We will call this variable the EARLY RESPONSE. The second dependent variable was the final evaluation of truth or falsity once all relevant information was available (i.e., at the third or fourth box). We will call this the LATE RESPONSE. Obviously, we only had data for the second dependent variable when the early response was to continue.

Factorial design of Experiment II Experiment II involved a 4*2 factorial design, just like Experiment I, the two factors being SENTENCE TYPE and CANONICAL. The four levels of the factor sentence type were identical to Experiment I. The

CANONICAL factor has two levels: [+CAN] and [–CAN]. In the [+CAN] condition the third or fourth box reveals the information that the canonical inference triggered by the target sentence is true, e.g., the third box contained the information that Max mixed a cocktail. As opposed to this, the [–CAN] condition reveals the information either in the third or fourth box that the canonical inference is false. For our example, Max did something other than mixing a cocktail (see Table 1.2).

Dependent variables of Experiment II As opposed to Experiment I, in Experiment II the second box which was uncovered always entailed that the exhaustivity inference triggered by the target sentence was false (hence "falsifier"). Accordingly, for our example in Table 1.2 above, the second box explicitly reveals that one of Tom, Jens, or Ben (in our example it is Ben) mixed a cocktail. Again, the first box never contained any information that would be relevant to the canonical or exhaustivity inference. With this background, we measured exactly the same two dependent variables as in Experiment I, i.e., EARLY RESPONSE and LATE RESPONSE. Of course, the evaluation of these dependent variables is radically different from Experiment I given the different information in boxes 2-4.

Fillers As filler items, we had sentences with the universal quantifier *jeder* 'everybody,' as in (15); expletive expressions beginning with *es ist klar* . . . , as in (16); subjects containing two conjoined proper nouns, as in (17); as well as the scalar expression *weniger als* 'fewer than,' as in (18). There were 8 lexicalizations per sentence type, randomized during presentation, and each participant heard the same 32 filler sentences. For the filler trials, the distribution of possible responses, i.e., verifiers of the canonical meaning and falsifiers of exhaustivity, was balanced across the four boxes with respect to the target stimuli. On top of deflecting participants' attention from the target constructions at issue, the fillers served the overall purpose of quality control in measuring the reliability of the experimental method.

(15) Jeder hat ein Buch ausgeliehen.
everyone has a book borrowed.
'*Everyone borrowed a book.*' UNIVERSAL

(16) Es ist klar, dass Ben eine Geschichte erfunden hat.
it is clear that Ben a story invented has
'*It is clear that Ben invented a story.*' EXPLETIVE

(17) Ben und Max haben einen Fehler korrigiert.
Ben and Max have a mistake corrected
'*Ben and Max corrected a mistake.*' CONJUNCTION

(18) Weniger als drei Leute haben ein Bankkonto eröffnet.
fewer than three people have a bank.account opened
'*Fewer than three people opened a bank account.*' SCALAR

In all targets and filler items, the verb that described the activity was in the present perfect in German, which in English is often translated as simple past, as in the glosses here.

1.3.2 Summary of the logic of the experiments

In both experiments, we measure whether and at which point the participants decide that the target stimuli are true or false given incremental evidence. Specifically, in both experiments we are interested in how participants will respond at Box 2, which, crucially, is where the two experiments differ: Experiment I verifies the canonical inference at Box 2, whereas Experiment II falsifies the exhaustivity inference at Box 2. The specific questions associated with the Early Response variable for the two experiments are as follows:

Experiment I attempts to establish whether for a cleft or for any of the other analyzed sentence types the knowledge that the canonical inference is true suffices to decide that the cleft sentence is true simpliciter, or whether the exhaustivity inference is also considered by participants. Clearly, if a participant chooses to give a 'true' judgment at this early evaluation stage, this means that for this participant the exhaustivity inference does not matter (enough) to justify further investigation (i.e., *non-exhaustive* responses). As opposed to this, if a participant decides to continue after Box 2, it means that the exhaustivity inference is important enough to be checked against the upcoming incremental information (i.e., *exhaustive* responses), and hence, we expect that the participant will answer 'true' in the final evaluation in the [+EXH] condition and 'false' in the [–EXH] condition . Experiment II attempts to establish whether or not knowing that the exhaustivity inference is false at Box 2 will suffice for participants to judge the whole sentence as false (i.e., *exhaustive* responses), or whether the participants consider it possible for the sentence to still be judged true (i.e., *non-exhaustive* responses). Clearly, if a participant chooses to continue after Box 2, the only rational reason to do so is that for this participant the canonical inference is sufficient to assign the value true. Therefore, we expect that the late evaluation answer only depends on the canonical factor, and we expect that in the final evaluation the participant will answer 'true' in the [+CAN] condition and 'false' in the [–CAN] condition.

Theoretical predictions The theoretical predictions for plain intonational foci and exclusives are identical on any of the major theories discussed above and will be

discussed together. Plain focus only gives rise to a weak pragmatic exhaustivity implicature, whereas exclusives give rise to a strong semantic and at-issue exhaustivity inference. Since we expect the exhaustivity inference to be frequently disregarded in the former and to be robustly present in the latter, the exhaustiveness patterns in the focus condition provide a baseline for *non-exhaustive* responses, and the exclusive condition provide a baseline for *exhaustive* responses. In terms of concrete experimental outcomes, this amounts to the following predictions for the early responses.

(19) Focus (implicatures calculated on the focus alternatives are not truth-functional)

Exp. I (verifier): Participants will give an early 'true' judgment at Box 2 since the exhaustivity implicature will not require further checking of the context to evaluate the asserted content.
Early Response ⇒ 'true'

Exp. II (falsifier): Participants will continue at Box 2 to evaluate the asserted content of the focus clause, since the exhaustivity implicature may not arise or may be cancelled if it is evidently false.
Early Response ⇒ 'continue'

(20) Exclusives (asserted exhaustivity is truth-functional)

Exp. I (verifier): Participants will continue until they have uncovered all the boxes to evaluate the asserted exhaustivity.
Early Response ⇒ 'continue'

Exp. II (falsifier): Participants will give an early false judgment at Box 2, since a sentence cannot be true if the semantic inferences it gives rise to are false.
Early Response ⇒ 'false'

For clefts and definite pseudoclefts, the predictions of the various theories are, naturally, different. Based on the discussion above we can reproduce the predictions from Table 4.1 (repeated below) in Table 1.3 including specific experimental outcomes associated with each type of theory.

While the translation of Table 4.1 to the experimental predictions in Table 1.3 is for the most part straightforward—that is, [–strength] exhaustivity can generally be expected to show parallel response patterns to plain intonational focus, and [+strength] exhaustivity to show parallel response patterns to exclusives—not all cells are subject to clear predictions. For instance, for the (A) pragmatic and (C) semantic IT-construction approaches, a non-parallel behavior of clefts and definite pseudoclefts in these particular environments is not necessarily predicted. After all, these approaches do not make any specific claims about definite pseudoclefts (the

Theory	± *strength*	± *parallel def. pse.*	*Exp. (Box 2)*	*Early Response at Box 2* *Clefts*	*Def. Pse.*
(A) *pragmatic*	–	–	I (verifier) II (falsifier)	**true** **continue**	~ continue ~ false
(B) *sem. def.*	+	+	I (verifier) II (falsifier)	**continue** **false**	continue false
(C) *sem. IT*	+	+/–	I (verifier) II (falsifier)	**continue** or **true** **false**	~ continue ~ false

Table 1.3 *Theoretical predictions for the early responses based on parameters of evaluation from Table 4.1. The symbol ~ indicates possible responses, since these approaches do not make any specific claims about definite pseudoclefts.*

~ symbol merely indicates possible responses in Table 1.3), for which reason it is certainly compatible with theories (A) and (C) if clefts and definite pseudoclefts were to pattern alike in these particular circumstances. Moreover, theory (C) allows for the possibility that the [+strength] exhaustiveness inference of clefts, which they consider semantic but not-at-issue, will be disregarded in Experiment I because of its being not-at-issue; therefore, in Experiment I theory (C) is compatible with both 'continue' as well as 'true' responses at Box 2. In a sense then, except for the (B) semantic definite theory—which has a clear position on each of the slots—theories (A) and (C) are less directly tested by our design. Be that as it may, as will become obvious in the following sections, the experimental results we obtained go beyond the predictions of any of these three theories.

1.3.3 Results

Experiment I Data preparation: For Experiment I, there was 1/1024 potential judgments at Box 1 for the target items, which was treated as an error and removed from the statistical analysis, since there is no discernible reason at this point in the procedure to make a truth-value judgment.

We start by describing the results of the Early Response. Exclusives elicited a judgment at Box 2 only 1% of the time (2/256 responses): Almost all participants chose to continue uncovering Box 3 and Box 4. Plain focus, by contrast, elicited a high percentage of judgments at Box 2, namely 74% of the time (189/256 responses): In the majority of trials participants made a 'true' judgment without checking the contexts to see whether exhaustivity held. As compared to the two control conditions, clefts and definite pseudoclefts fell somewhere in the middle—at least

in the overall numbers and proportions across all participants, but see the post hoc analysis discussed in Section 1.3.4—with clefts eliciting a judgment 43% (110/255 responses), and definite pseudoclefts 41% of the time (105/256 responses). Since in all cases in which the Early Response was a judgment it was a 'true' judgment, we do not treat this as a three-valued parameter but as a two-valued parameter, i.e., whether a judgment happened. See Figure 1.2 for the observed proportions of Early Responses in Experiment I (left graph, triangles) made per sentence type.

We conducted a generalized linear mixed effects model for binomial data to compare statistically the likelihood of participants making a ('true') judgment.[10] We used treatment contrasts encoded as numeric covariates for the SENTENCE TYPE condition, in which clefts were compared to each of the other levels. Crucially, no significant difference was found between clefts and definite pseudoclefts ($\hat{\beta}$ = –0.2831, SE = 0.3076, z = –0.920, p = 0.357); by contrast, focus was significantly more likely to elicit 'true' judgments ($\hat{\beta}$ = 4.1125, SE = 0.9120, z = 4.510, p = 6.5e–06). Note that given the difficulty of making meaningful comparisons with percentages close to zero, the exclusive condition was not included in the generalized linear mixed model for Experiment I. See Figure 1.2 for the back-transformed model-predicted proportions (left graph, dots with 95% confidence intervals) for Early Responses. As can be seen by the differences between the observed and model-predicted proportions—most notably in the plain focus conditions in both experiments, in which the observed proportions lie outside of the 95% confidence intervals—the model predictions very poorly match the observed data. This shows that the mixed-effects logistic regression is an inappropriate model for the data. However, once the participants are divided into responder groups based on the response patterns for clefts as in the post-hoc exploratory analysis presented below, the model predictions do match the data.

In the cases when participants chose to continue (Late Response): In the [+EXH] condition for Box 3 or Box 4, in which exhaustivity was not violated, the final judgment was consistently 'true'; by contrast, in the [–EXH] condition, in which exhaustivity was violated, the final judgment was consistently 'false.' This is shown in Table 1.4.

10 We ran generalized linear mixed-effects models for binomial data in R (GPL-2|GPL-3: v.3.3.3, R Core Team 2017) with the *lme4* library (GPL-2|GPL-3: v.1.1-15, Bates, Mächler, et al. 2015). Following recommendations in Bates, Kliegl, et al. 2015, and utilizing the *rePCA* function in the *RePsychLing* library (MIT, v.0.0.4) (available at https://github.com/dmbates/RePsychLing), we included variance components/correlation parameters in the random-effects structure supported by the data. The resulting parsimonious models were as follows. Experiment I: $glmer(TVJ.Box2 \sim DefPse + Foc + (1 + Foc \,||\, Participant) + (1 + Foc \,||\, Item), family = binomial)$. Experiment II: $glmer(TVJ.Box2 \sim DefPse + Excl + Foc + (1 + DefPse + Excl + Foc \,||\, Participant) + (1 + Foc \,||\, Item), family = binomial)$. We ignored the late exhaustivity factor in our computation, because in this early stage of evaluation it plays no role.

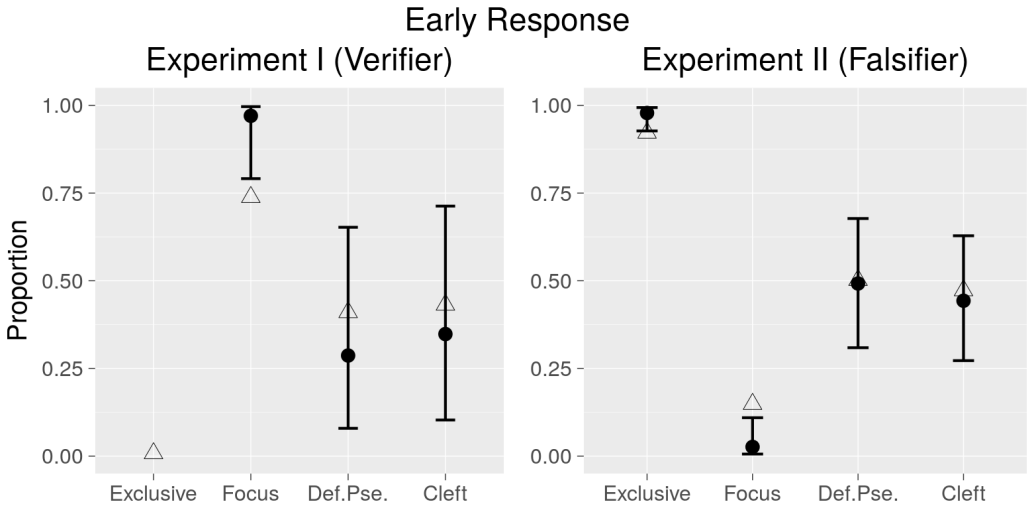


Figure 1.2 *Observed (triangles) and back-transformed predicted proportions (dots, with 95% confidence intervals) for Early Responses for Experiment I (left) and Experiment II (right): judgment = 1, continue = 0. Given percentages close to zero for the exclusive condition in Experiment I, only the observed proportions are presented.*

			Exclusive	Focus	Def.Pse.	Cleft
Exp. I (verifier)	[+EXH]	'true'	98% (123/126)	100% (32/32)	99% (76/77)	99% (71/72)
	[–EXH]	'false'	99% (127/128)	91% (30/35)	91% (67/74)	93% (68/73)
Exp. II (falsifier)	[+CAN]	'true'	14% (1/7)	96% (105/109)	85% (53/62)	87% (58/67)
	[–CAN]	'false'	92% (12/13)	99% (108/109)	100% (65/65)	97% (65/67)

Table 1.4 *Late responses as percentages (fractions in parentheses) in [+/–*EXH*] conditions in Experiment I and [+/–*CAN*] conditions in Experiment II.*

Experiment II Data preparation: for Experiment II, there were 3/1024 'true' judgments at Box 2 upon encountering a falsifier of exhaustivity, which were treated as errors and removed from the statistical analysis since there is no logical reason to make a 'true' judgment given the information revealed. Again, since in all cases in which the Early Response was a judgment it was a 'false' judgment, we treat this as a two-valued parameter, i.e., whether a judgment happened.

Exclusives elicited 'false' judgments 92% of the time at Box 2 (236/256 responses): Most participants chose not to continue uncovering further boxes. By contrast, plain focus elicited 'false' judgments only 15% of the time (38/256 responses), with most participants choosing to continue. Definite pseudoclefts elicited 'false' judgments 50% of the time (128/255 responses), and clefts were very similar in eliciting judgments 47% of the time (120/254 responses). See Figure 1.2 for the observed proportions of Early Responses in Experiment II (right graph, triangles) made per sentence type.

We conducted a generalized linear mixed-effects model for binomial data to compare the likelihood of participants making a ('false') judgment. Again, we used treatment contrasts encoded as numeric covariates: Clefts were the baseline comparison for all other sentence types. In both experiments, there was no significant difference found between clefts and definite pseudoclefts ($\hat{\beta}$ = 0.1978, SE = 0.2527, z = 0.782, p = 0.434). By contrast, exclusives were significantly more likely to elicit 'false' judgments ($\hat{\beta}$ = 4.0413, SE = 0.5907, z = 6.842, p = 7.81e–12), while focus was significantly more likely to elicit 'continue' ($\hat{\beta}$ = –3.3849, SE = 0.7151, z = –4.733, p = 2.21e–06). See Figure 1.2 for the back-transformed model-predicted proportions (right graph, dots with 95% confidence intervals) for Early Responses. Note again the differences between the observed and model-predicted proportions, namely in the focus condition, showing that the predictions from the model very poorly match the observed data; however, once participants are divided into responder groups as in the post-hoc analysis below, there is in fact a better match between the model predictions and the data.

In the cases when participants continued uncovering (Late Response), the final judgment in the [+CAN] condition was consistently 'true' (with the exception of exclusives in Exp. II; however, note the very low number of data points—i.e., most participants had made an early judgment), whereas the final judgment in the [–CAN] condition was consistently 'false.' This is shown in Table 1.4.

Post hoc analysis In both experiments the ratio of continue and true/false judgments for clefts and definite pseudoclefts as an early response were about 50-50, instead of the predicted 0-100 or 100-0 (modulo noise). A natural question is, then, whether the midway average arises due to differences in participant's behavior or whether the items created the variation. In a post hoc analysis, we found that when analyzing participant behavior individually two main groups emerged for clefts: In both experiments participants treated clefts either as exhaustively as they did exclusives (Experiment I: 19 participants; Experiment II: 14 participants) or as non-exhaustively as they did plain focus (Experiment I: 13 participants; Experiment II: 16 participants). Only two participants across both experiments responded at chance levels (Experiment II: 2 participants).

These categories were based on percentages for the response patterns in clefts, since after data preparation (in which erroneous judgments were removed; see above) not all participants had the same denominator for total possible judgments at Box 2. The two categories were calculated as follows. Participants who chose 'true' for clefts 60% or more of the time fell into the non-exhaustive interpretation group, generally treating clefts more like focus (i.e., they made a 'true' judgment upon verifying the canonical meaning of the sentence); and participants who in Experiment I chose 'true' for clefts 40% or less fell into the exhaustive interpretation group, treating clefts more like exclusives by continuing a majority of the time. Conversely, in Experiment II if participants made a 'false' judgment for clefts 60% or more of the time, they fell into the exhaustive interpretation group, treating the clefts as they did exclusives; and if they made a judgment 40% or less of the time, they were in the non-exhaustive interpretation group (i.e., they generally chose 'continue' upon falsifying exhaustivity, similar to focus). In both experiments, if participants made a judgment between 40–60% of the time, they fell into the chance group. Observed proportions (triangles) for each group are shown in Figure 1.3 for Experiment I and in Figure 1.4 for Experiment II.[11]

11 The late responses of those who continued, i.e., the exhaustive group in Experiment I and the non-exhaustive group in Experiment II, support our choice of labels for the two groups. The participants of the exhaustive group did actually judge false if exhaustivity was violated in Box 3 or 4 for Experiment I, and the participants of the non-exhaustive group in Experiment II judge true when the canonical was true in Box 3 or 4 even though exhaustivity was violated in Box 2 (see Table 1.4).

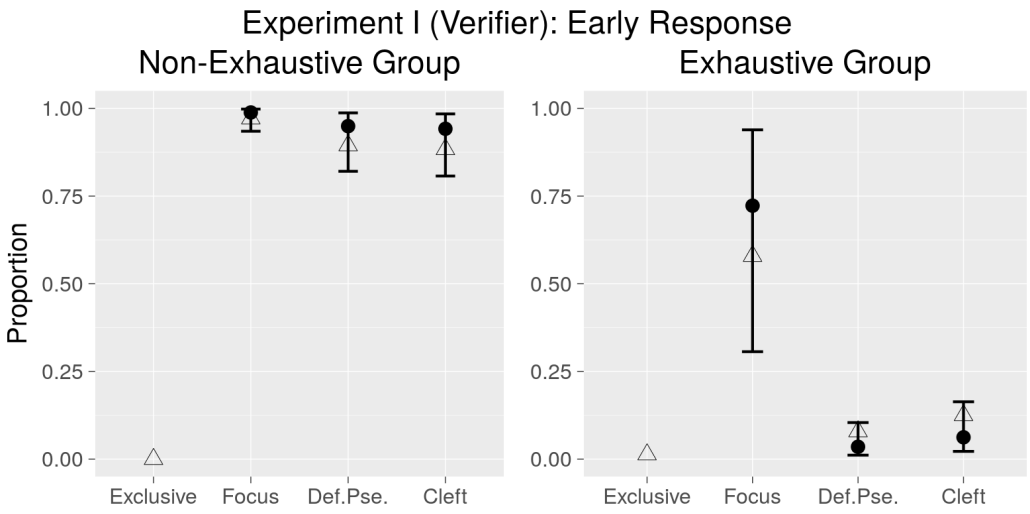


Figure 1.3 *Observed (triangles) and back-transformed predicted proportions (dots, with 95% confidence intervals) for Early Responses for non-exhaustive (left) and exhaustive (right) groups: judgment = 1, continue = 0. Given percentages close to zero for the exclusive condition in Experiment I, only the observed proportions are presented.*

The results presented in Figures 1.3 and 1.4 show that participants who treated clefts more like exclusives also treated definite pseudoclefts more like exclusives, and the same pattern was found for those who treated clefts like focus. Note again that the two experiments were run with different participants and these findings do not suggest that one and the same participant behaves in an erratic way across the experiments. On subsets of the data corresponding to the exhaustive and non-exhaustive groups for each experiment, we conducted a generalized linear mixed-effects model for binomial data to test the likelihood of making a judgment. We wanted to see in the non-exhaustive groups, whether clefts differed from focus, and in the exhaustive groups, whether clefts differed from exclusives.

For the non-exhaustive groups (left graphs in Figures 1.3–1.4) there was a significant difference found between clefts and focus in Experiment I ($\hat{\beta}$ = 1.6571, SE = 0.6992, z = 2.370, p = 0.0178), with the focus condition more likely to elicit 'true' judgments in comparison to clefts; but by contrast there was no significant difference found between these two sentence types in Experiment II ($\hat{\beta}$ = –2.0136, SE = 1.1225, z = –1.794, p = 0.0728). Inversely, for the exhaustive groups (right graphs in Figures 1.3–1.4) , a significant difference was found between clefts and exclusives in Experiment II ($\hat{\beta}$ = 1.2883, SE = 0.4970, z = 2.592, p = 0.00953), with

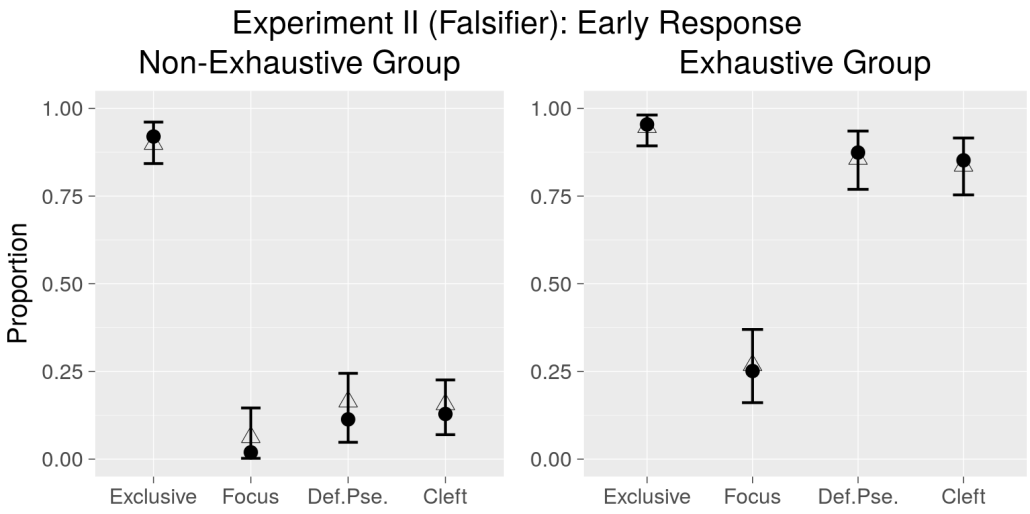


Figure 1.4 *Observed (triangles) and back-transformed predicted proportions (dots, with 95% confidence intervals) for Early Responses for non-exhaustive (left) and exhaustive (right) groups: judgment = 1, continue = 0.*

exclusives more likely to elicit 'false' judgments than clefts.[12] Note again that given the difficulty of making meaningful comparisons with percentages very close to zero, the exclusive condition was not included in the generalized linear mixed model for Experiment I. Crucially, the post hoc analysis found no statistical difference between clefts and definite pseudoclefts in both groups of participants in both experiments (Exhaustive Group, Exp. I: $\hat{\beta}$ = –0.5993, SE = 0.4206, z = –1.425, p = 0.154, Exp. II: $\hat{\beta}$ = 0.1882, SE = 0.4156, z = 0.453, p = 0.65071; Non-Exhaustive Group, Exp. I: $\hat{\beta}$ = 0.1479, SE = 0.4879, z = 0.303, p = 0.7618, Exp. II: $\hat{\beta}$ = –0.1447, SE = 0.4965, z = –0.292, p = 0.7707), albeit in two different ways.

12 An anonymous reviewer pointed out that such results could also point toward a three-way distinction between exhaustivity (with exclusives), partial exhaustivity (with clefts and pseudoclefts), and non-exhaustivity (with plain intonational focus), perhaps in parallel to the distinction between factive, semi-factive, and non-factive predicates (Karttunen 1971). However, an account along these lines would have nothing to say on the exact source of partial exhaustivity in clefts and pseudoclefts. In light of this, we favor an account in which the observed differences between clefts/pseudoclefts and focus (non-exhaustive group) or exclusives (exhaustive group) are accounted for on the basis of different interpretive processes underlying the exhaustivity inferences in exclusives (truth-functional entailments), clefts/pseudoclefts (accomodation of implicit discourse antecedent), and plain focus (scalar implicature), respectively. In section 1.4, following the approach in Pollard & Yasavul (2016), we will propose a pragmatic analysis of cleft/pseudocleft exhaustivity that sets it apart from the semantically entailed exhaustivity of exclusives, on the one hand, and the focus-driven scalar exhaustivity implicature of plain focus, on the other.

1.3.4 Discussion

In this section we revisit the logic of the experiments and theoretical predictions discussed in Section 1.3.2. First, we evaluate participant response patterns in terms of the logic of the experiments, discussing how the results show that the participants understood the task and acted accordingly. This will allow us to disregard from further discussion the late evaluation results, for given the logic of our experiments they are predictable from the early evaluation data. In the second step, we discuss how the results relate to the theoretical predictions. In doing the latter, we will also discuss the results of the post hoc analysis, which have somewhat surprising consequences that are unexpected in light of the existing theoretical literature.

Evaluation of the logic of the experiments In both experiments, we measured whether and at which point the participants decided that the target stimuli were true or false given the incremental evidence provided. Of particular interest was Box 2, which verified the canonical inference in Experiment I and falsified exhaustivity in Experiment II. The primary questions for the early responses and expected response patterns for the late responses were as follows.

- Experiment I – Early Response (Box 2): Was it enough to verify the canonical inference to make a truth-value judgment (*non-exhaustive* response), or was exhaustivity also considered (*exhaustive* response)? Late Response (Box 3/4): In the latter case, i.e., for those participants for whom exhaustivity was important enough to continue uncovering, we expect 'true' responses in the [+EXH] condition and 'false' responses in the [–EXH] condition.
- Experiment II – Early Response (Box 2): Does falsifiying exhaustivity suffice to judge the whole sentence as false (*exhaustive* response), or did participants consider it still possible to judge the sentence as true by continuing to uncover boxes (*non-exhaustive* response)? Late Response (Box 3/4): In the latter case, i.e., for those participants for whom violating exhaustivity was not sufficient to make a 'false' judgment, we expect 'true' responses in the [+CAN] condition and 'false' responses in the [–CAN] condition.

Indeed, when participants continued uncovering Box 3 and Box 4, the expected patterns for the late responses were precisely what we found (see Table 1.4 on page 49), and hence the late evaluation data substantiate that participants understood the logic of the experiment; beyond this, the late evaluation data are of no interest.

In a simple semantic model in which we have a sentence E licensing two inferences p and q, it should be obvious that experiments manipulating verification/falsification in the way reported above are expected to produce mirror image

results. If a hypothetical speaker finds that *p* is true and decides that this information suffices to judge *E* as true, she considers that the inference *q* is, in some sense, irrelevant; accordingly, the speaker can be expected not to judge *E* as false if she was presented with the evidence that *q* is false. For such a speaker, sentence *E* simply means *p*, whereas *q* is not strictly entailed and therefore neglectable. Conversely, if a hypothetical speaker finds that *p* is true and yet decides to check the truth of *q* in order to evaluate whether *E* is true, this person is expected to judge *E* as false once she sees that *q* is false. For such a speaker *E* means at least the logical conjunction between *p* and *q*. (For the sake of clarity, recall that in our experiments the participants in Experiment I were distinct from participants in Experiment II.)

Assuming that exhaustivity is an inference of some type acknowledged in the literature for all the sentence types tested in the experiment, it follows that both an early 'true' judgment in Experiment I, and an early continue decision corresponding to a late 'true' judgment in Experiment II, will indicate that the exhaustivity inference is not as strong as a semantic inference would be expected to be.

Even though the two experiments are mirror images of each other on a simple model, conducting both Experiment I and Experiment II in tandem is not superfluous. Consider the possibility that *p* is an at-issue inference of *E* whereas *q* is a not-at-issue inference. The literature, following Tonhauser et al. 2013, seems by and large to converge in acknowledging that not-at-issue inferences do not form a homogeneous group. A conceivable class of not-at-issue inferences might be such that, on hearing *E*, the not-at-issue inference *q* simply does not come to mind as something that has been conveyed. Call this class *not-immediate* inferences and leave it open whether this is an empty class. Crucially, a hypothetical speaker could simply judge *E* as true without checking for a not-immediate inference *q*, precisely because *q* did not come to mind, but was potentially taken for granted or forgotten altogether. As opposed to this, when faced with the explicit falsity of *q*, a hypothetical speaker is no longer in a state of mind in which *q* can be disregarded. Hence, in this case, we would expect *E* to be judged largely true when verified and false when falsified. As the experimental results show, this was not the case, however: The exhaustivity inference of clefts did not behave like not-immediate inferences would be expected to behave.

Moreover, apart from eliminating the above-mentioned source of confound, including both verification and falsification in the experimental setup also serves the purpose to detect and overcome potential biases of participants toward judging sentences true rather than false. If such were the case, we would expect that participants judge the exhaustivity inference true more often in Experiment I than false in Experiment II. The results clearly show that this did not happen.

Evaluation of the theoretical predictions The results show that clefts and definite pseudoclefts behave in an unexpected way when compared to theories (A) to (C)

from Table 1.3. In particular, in both experiments the ratio of continue and true/false judgments as an early response were about 50-50, instead of the predicted 0-100 or 100-0 (modulo noise). Moreover, clefts and definite pseudoclefts show neither a similarity to exclusives nor a similarity to plain focus.

More interestingly, in the post hoc analysis it was found that participants fell into two groups, and about half of the participants acted as the semantic definite account would have it: These participants judged definite pseudoclefts and clefts almost as exhaustively as exclusives—that is, they cared about exhaustivity and behaved accordingly. By contrast, the other half of participants showed the exact opposite behavior—these participants were willing to identify the referent x in a way which was not exhaustive with respect to P. This constitutes a serious puzzle for the semantic definite account, as one would not expect a semantically hardwired inference to be available for only half of the population. At the same time, it is a serious problem for the pragmatic approach as well, since the exhaustive group did not interpret plain focus in a parallel way to clefts.

In light of this, it is implausible to assume that, for the exhaustive group, the exhaustivity inference in clefts is an implicature that happened to remain uncancelled, while it is subject to cancellation with the non-exhaustive group. More generally, the different behavior of plain focus vs. clefts suggests that the exclusion of salient focus alternatives, possibly per implicature, is not the driving force behind the exhaustivity inference in the latter. Finally, given that exhaustivity is a significant inference in communication, it will not do to assume that there are two dialects of German in order to explain the observable differences between participants. If there were such two dialects (that were not geographically separated by a natural border), their speakers would be expected to show a systematic failure of mutual understanding when a cleft or pseudocleft is used. Instead, a valid explanation for the observed pattern should rather involve some parameter of evaluation that can be reasonably taken to differ for the two participant groups, such that the exhaustivity inference is present only if that parameter is set to a certain value.[13] Except for Pollard & Yasavul 2016, none of the above mentioned accounts involve such a parameter.

In conclusion, no difference between clefts and definite pseudoclefts was found in the two experiments conducted. Critically, both sentence types lacked an exhaustive interpretation with about half of the participants. The exhaustivity inference in clefts and definite pseudoclefts was not found to be strong: It was neither robust nor

13 Moreover, recall that the domain of quantification was explicitly fixed in the experiments, and there was also no scalar ordering of alternatives in terms of noteworthiness or unlikelihood. These features of the experimental set-up rule out typical attempts at explaining exhaustivity violations away as only apparent (É. Kiss 2010, Skopeteas & Fanselow 2011). The explanation for the split behavior of participants in the cleft and definite pseudocleft conditions must lie elsewhere.

systematic. This observation is incompatible with any of the three main theoretical approaches to cleft exhaustivity in the literature, and calls for an alternative account.

1.4 Analysis

In this section, we present our main analysis based on the experimental observations above. In particular, we claim that participants only got an exhaustive reading if they took clefts and definite pseudoclefts to anaphorically refer to an implicit question (following Pollard & Yasavul 2016). We present the core of the analysis in Section 1.4.1. In Section 1.4.2, we discuss how our proposal for clefts extends to the case of German definite pseudoclefts with *derjenige*.

1.4.1 Anaphoricity of clefts

It is standardly assumed in the literature that clefts have anaphoric potential (Prince 1978, Horn 1981, Soames 1989, Delin 1992, Hedberg 2000, and many others): They introduce as part of their constructional meaning a presupposition that marks the information conveyed by the cleft as known-fact (Prince 1978) or, simply, as anaphoric (Delin 1992); see in particular Delin 1992 for ample empirical evidence for the anaphoricity of clefts. The anaphoric potential of clefts can be formally expressed in the form of an existence presupposition, following van der Sandt 1989 and Rooth 1996. We further assume that existential presuppositions must be licensed in discourse. Following van der Sandt (1992) and many others, we adopt a dynamic model of discourse and assume that there are two main options for licensing presuppositions, viz. accommodation and binding. Moreover, we follow van der Sandt (1992) and many others in assuming that, whenever possible, presupposition binding in context must be chosen over accommodation.

Empirically, this assumption amounts to the observation that whenever a cleft follows an explicit discourse referent with the relevant properties in the preceding discourse, the cleft must by necessity refer back to that discourse referent. This is shown in (21), which—on its only licit interpretation—suggests that the discourse referent introduced by the indefinite in the first sentence is anaphorically picked up and further specified in the cleft sentence. In other words, the existential presupposition of the cleft is necessarily dynamically bound by material in the first sentence.

(21) A: Judy was looking for somebody all afternoon.
B: It was her youngest daughter that Judy was looking for.

Turning to the experimental setting of our experiments, there is no linguistic context against which to evaluate the audio stimulus: As a result, we assume that the existential presupposition of the cleft condition must be accommodated.[14] This amounts to saying that the hearer will integrate into her discourse model some discourse referent with the relevant property described by the cleft relative that she takes the experimental speaker to (anaphorically) refer to.

Crucially, we do not adopt claims in Szabolcsi 1994 (on pre-verbal focus in Hungarian) and Percus 1997 (on English *it*-clefts) that the existential presupposition of cleft sentences comes with an obligatory maximality effect, say a maximality presupposition built into the structure of clefts. Our reasons for rejecting this common assumption for clefts are as follows: If paired with some sort of maximality effect, the existential presupposition will require the discourse to contain a bound or accommodated discourse referent x, such that x has the property P described in the cleft relative, and nobody other than x (in the relevant domain) has property P. Given this, there are two possibilities to consider.

The first possibility is to assume that the compositional semantics of clefts is built around an identity statement such that the cleft pivot x equals the discourse referent y described by the cleft relative ($x = y$). In this case, the presupposed maximal discourse referent with cleft relative property P, namely y, will be identical to the cleft pivot x. This in turn amounts to clefts being semantically exhaustive, in contradiction to our experimental findings. So, we must reject this possibility. The second possibility is to assume that the compositional semantics of clefts does not involve an identity relation, but a plain predication relation instead ($P(x)$). In this case, the cleft would presuppose there to be a maximal discourse referent x with cleft relative property P, and in addition it would assert that it is the cleft

14 An anonymous reviewer asked how we can be sure that the existence presupposition was accommodated rather than simply ignored. Indeed, in the experimental literature there are examples of participants outright ignoring presuppositions in unembedded environments, such as, for instance, with the German iterative *wieder* 'again' (Tiemann 2014). In order to account for such cases, Tiemann proposed a maxim of interpretation called *Minimize Accommodation* (MA)—the only principled account for such data the authors are aware of—which dictates: "Do not accommodate a presupposition unless missing accommodation will lead to uninterpretability of the assertion!" (43). As Tiemann (2014: 44) writes: "this is a principle that *every interpreter* adheres to when faced with a situation in which s/he cannot ask for further information regarding the PSP" [emphasis added]. Assuming MA is even applicable here, we think such a maxim makes the wrong predictions in light of our data: Most or all participants would be expected to leave the existence presupposition unaccommodated in our experiment (cf. the ignored presupposition of *wieder* in Tiemann 2014). If that were the case, however, then clefts and focus would end up having the exact same semantic contribution and would be predicted to elicit identical response patterns, contrary to what we found; furthermore, we would have no satisfactory account for why half the population treated clefts as exhaustively as exclusives. Thus, we rather assume that participants accommodate the anaphoric existence presupposition, from which one can derive the exhaustive/non-exhaustive interpretation.

pivot x that has the property P. It is easy to see, as pointed out by Büring & Križ (2013), that maximality is vacuously satisfied in this scenario whenever the existence presupposition is satisfied. The reason is that mere existence already entails the existence of a maximal witness, such that the presence of an appropriate referent in the discourse will automatically satisfy maximality. So, postulating an additional maximality presupposition in clefts will either come out as empirically false or as semantically vacuous, depending on the compositional analysis of clefts chosen.

Observe that, up to this point, our analysis shares the fate of Horn's (1981). Horn also assumed that clefts come with an existential presupposition, but on top of this he was forced to invoke a general pragmatic principle in the form of a generalized conversational implicature in order to derive the exhaustivity inference; see section 1.2.1 for discussion. Again, the assumption of a general pragmatic principle cannot account for the experimental data at hand, as it would predict a uniform behavior of participants in the experiment, contrary to fact. Instead we propose that part of what the experiment participants did was to reason about the anaphoric antecedent of the cleft's existential presupposition. Building on an idea in Pollard & Yasavul 2016, we will argue that there are two such reasoning procedures, resulting in an exhaustive or non-exhaustive interpretation, respectively. Importantly, both procedures are compatible with an underlying identificational semantics of clefts, in which the value of a variable x is equated with the denotation of the focused cleft pivot (see below). The relevant question is how the value for the variable is resolved to some salient discourse antecedent.

According to Pollard & Yasavul (2016), one way of constructing a suitable discourse referent x in the absence of explicit context consists in taking the cleft to answer an implicit *wh*-question. That is, participants may take a cleft of the form "It is α who P" to address the question issue "who P?", thus resolving the existence presupposition to a maximal discourse referent x with property P. Linking this with an identificational at-issue semantics for clefts, namely $x = \alpha$, the result will be that the maximal individual x with property P equals the pivot α, which comes down to an exhaustivity claim. This account of cleft exhaustivity relies on the assumption first made by Hamblin 1957 that questions invariably denote sets of *complete* answers, the cleft serving to identify one of those complete answers. The second strategy, according to Pollard & Yasavul (2016), consists in accommodating a non-maximal discourse referent, as is the case, e.g., with indefinite antecedents; see our example (21). On this resolution of the discourse antecedent, the cleft simply expresses that there is some x with property P, and $x = \alpha$, which does not trigger an exhaustivity inference. However, given that indefinites have also been associated with (potential) questions in recent inquisitive semantic analyses (e.g., Onea 2016), it is not obvious to us whether the two resolution strategies should be tied to the presence or absence of a context question.

In view of this problem, we propose the following modified account of the behavior of the exhaustive and non-exhaustive groups in our two experiments, which retains the central insight of Pollard & Yasavul (2016). Members of the exhaustive group predominantly accommodate a discourse antecedent that is maximal with respect to the backgrounded property *P*, viz. (22a). When casting this in a question-based discourse analysis (Roberts 2012), the corresponding QUD could be either an exhaustively interpreted *wh*-question, or else an identification question (22b). The discourse referent *x* can be modeled with the iota-operator, and the meaning of the exhaustive-interpreted cleft is shown in (22c).

(22) a. There is a maximal (sum) individual *x* that mixed a cocktail.
It's MAX that mixed a cocktail.
b. $\text{Who}_{\text{COMPL}}$ mixed a cocktail? / Who is the maximal *x* that mixed a cocktail?
c. ASS: *x* = max, PSP: ∃x[x mixed a cocktail]
⇒ ιx[x mixed a cocktail] = max

Members of the non-exhaustive group, by contrast, predominantly chose to accommodate an indefinite (non-maximal) discourse antecedent, viz. (23a), as in Pollard & Yasavul 2016. The indefinite gives rise to the *potential question* in (23b) (Onea 2016), an open complement question, resulting in a non-exhaustive interpretation. Technically, the non-maximal discourse referent *x* can be modeled by means of a choice function (Reinhart 1997, Winter 1997), which picks a random element from the backgrounded cleft property *P*, as in (23c).

(23) a. Somebody mixed a cocktail.
It's MAX that mixed a cocktail.
b. Who is this somebody that mixed a cocktail? / Who was it?
c. ASS: *x* = max, PSP: ∃x[x mixed a cocktail]
⇒ f(⟦mixed a cocktail⟧) = max

The foregoing assumptions suffice to explain our experimental findings. On the proposed analysis, the exhaustivity inference is a pragmatic effect that can be reliably predicted in explicit contexts, but which is not mandatory in the absence of overt linguistic context.[15] Depending on whether participants choose a maximal or

15 One reviewer pointed out that given the contextual sensitivity of the exhaustivity interpretation one might instead model the exhaustivity inference as a particularized conversational implicature (PCI). We cannot exclude that possibility, but how to go about spelling out the analysis is not obvious to us. The main problems as we see it are determining which context would need to be assumed to derive the PCI, and moreover, what role the existence presupposition would play, since it would be necessary for such an analysis to predict that canonical sentences do not give rise to the same exhaustivity inference in these contexts.

an indefinite (non-maximal) discourse antecedent, the cleft triggers an exhaustive or a non-exhaustive interpretation, respectively—responses in the early and late measures will pattern accordingly (see Section 1.3.4 under *Evaluation of the logic of the experiments*). Importantly, the source of the exhaustive effect does not lie the underlying identificational semantics of the cleft *per se*, but it lies in the different mechanisms for assigning a value to the variable x in the asserted identificational statement $x = max$, i.e. iota-operator vs. choice-function. Following Reeve's (2012) analysis of the pronoun *it* in *it*-clefts as a referring expression, the underlying identificational semantics of clefts is derived by equating the meaning of the cleft pivot with the meaning assigned to this pronoun, i.e., a contextually salient discourse antecedent. As the literal meaning of *it*-clefts no longer makes reference to maximality/uniqueness, there is no longer a tension between the fact that such sentences express an identificational statement and the fact that they can be non-exhasutive.[16] Finally, while we take exhaustive inferences with clefts to be pragmatic in nature (Horn 1981, 2014), on our account, they have nothing to do with the exhaustification of focus alternatives, nor with scalar implicatures computed over focus alternatives, pace De Veaugh-Geiss et al. 2015.

1.4.2 The case of definite pseudoclefts

What remains to be done is to show how the pragmatic analysis developed for clefts can be extended in order to capture the parallel interpretive properties of definite pseudoclefts in our experiments. As mentioned in section 1.2.1 definites do not seem to constitute a homogeneous class. In the following we want to argue against definite pseudoclefts falling into the same category as semantically unique definites. In particular, we claim that for definite pseudoclefts in German, deriving exhaustivity with an anaphoric familiarity analysis à la Heim 1982 better captures the results reported here.

Following a long list of scholars ranging from Frege (1892) to Coppock & Beaver (2015), definite descriptions in general are commonly treated as triggering a uniqueness presupposition as in (24).

(24) 'The NP_{sg}': Presupposes that the extension of NP has the cardinality smaller or equal to 1.

16 Alternatively, one could analyze *it*-clefts as structurally on a par with definite pseudoclefts (see Section 1.4.2), and assume, following Percus (1997), that both sentence types contain a (covert) strong, anaphoric definite determiner in the sense of Schwarz (2009). The individuals picked out by such determiner are unique in a weaker sense. They refer to the unique contextually salient discourse antecedent satisfying the backgrounded predicate P. As shown in the main text, such discourse antecedents can also be provided by indefinite NPs, resulting in non-exhaustive cleft interpretations.

This predicts a strong exhaustivity effect with definite pseudoclefts, contrary to what we found and reported here. In order to account for the observed absence of exhaustivity effects with about half of the participants, we instead need to resort to a familiarity-based analysis of definiteness.

More precisely, we would like to propose that definite pseudoclefts do indeed express anaphoric reference as part of their conventional meaning, as evidenced by their discourse-semantic behavior and by their morpholexical make-up. These two aspects distinguish definite pseudoclefts from regular definite descriptions which we will discuss in the following. Observe that definite peudoclefts are deviant as discourse openers, especially in comparison to their plain definite description counterparts, even if the two types of definite expressions have the same descriptive content. The relevant contrast is illustrated in (25). Example (25b) allows for easy accommodation of the fact that the lord, whoever that may be, has been murdered by someone, thereby triggering the interpretation that the gardener was the murderer. Example (25a), in contrast, resists such an interpretation. The most natural interpretation for (25a) is that it presupposes that the murder of the lord has already been the topic of discussion in the preceding discourse, either explicitly or implicitly. This being a condition on discourse structure, and not on the external world as such, it is rather hard to accommodate, especially at the beginning of a story.[17]

(25) Out Of The Blue

a. #Derjenige, der den Lord umgebracht hat, war der Gärtner.
the.one who the lord murdered has was the gardener
'*The one who murdered the lord was the gardener.*'

b. Der Mörder des Lords war der Gärtner.
'*The murderer of the lord was the gardener.*'

Likewise, (26a) is only acceptable if it is already evident in the preceding discourse that there is a man standing behind the hearer, or at least that there is a salient group of men with the speaker intending to refer to one of them. (26b), in contrast, would be a perfectly natural statement in a general discussion about hats, signaling by way of random example that the man behind the hearer has a relevant kind of hat. In such contexts, (26a) is not licit.

17 Arguably, the meaning of *derjenige*-phrases may be more complex than described here. Possibly, they also require explicitly or implicitly mentioned alternatives to the DP in the preceding discourse. Since those alternatives were always given in our experiment (in the form of the four roommates) this issue does not influence our analysis. Furthermore, an anonymous reviewer pointed out to us that the observed difference in (25) does not seem to exist in English. The difference between German and English might arise from the fact that German *derjenige*-phrases contain a demonstrative element *jene-* as opposed to the pro-NP form *one* in English *the one who*-phrases.

(26) a. #Derjenige, der hinter dir steht, hat einen tollen braunen Hut.
the.one who behind you stands has a fancy brown hat
'*The one who's standing behind you is wearing a fancy brown hat.*'
b. Der Mann hinter dir hat einen tollen braunen Hut.
'*The man behind you is wearing a fancy brown hat.*'

These observations about regular definite descriptions and definite pseudoclefts again speak in favor of them falling into different categories.

Having established that definite pseudoclefts express an anaphoric relationship in the form of an existence presupposition rather than uniqueness in the utterance situation, we can apply the same reasoning as for the cleft case, which gives us precisely the same predictions. To be concrete, we analyze the definite DP in definite pseudoclefts as a strong anaphoric determiner in the sense of Schwarz 2009, which evaluates uniqueness against some contextually salient discourse antecedent. This is shown in the following:

(27) ASS: ιy[y mixed a cocktail $\wedge$ $y = \mathbf{x}$] = max, PSP: $\exists$x[x mixed a cocktail]

As the individual denoted by the definite pseudocleft DP is no longer just evaluated relative to the background predicate *P*, we do not expect strong uniqueness or exhaustivity effects for this construction. Same as for *it*-clefts, depending on whether the variable *x* in (27) is resolved to a maximal/unique or to some indefinite antecedent, the pseudocleft will end up with an exhaustive or non-exhaustive interpretation, respectively. This accounts for the parallel behavior of clefts and definite pseudoclefts in our experiments, irrespective of whether or not the two sentence types share the same underlying syntax; see footnote 16.

1.5 Conclusion

In this paper, we have reported the results of two offline experiments on cleft exhaustivity in the incremental information-retrieval paradigm. It was shown that clefts and definite pseudoclefts are treated alike by the participants of a verification and a falsification experiment, in contrast to sentences with plain intonation foci and to sentences with exclusive particles. In particular, the exhaustivity inference in clefts and definite pseudoclefts is more pronounced than with plain focus, while being less strong than with exclusive particles.

We have argued that the non-systematic and non-robust nature of the exhaustivity effect is not accounted for by existing theoretical accounts, be they semantic or pragmatic. Moreover, a post hoc analysis further unveiled that participants showed systematic differences and parallel behavior in the interpretation of clefts and def-

inite pseudoclefts. About half of the participants treated both construction types consistently as exhaustive, while the other half treated both construction types as non-exhaustive. Again, this finding poses a challenge to semantic theories of cleft exhaustivity.

In response to these data, we argue that there must be some pragmatic component in the derivation of cleft exhaustivity. Our approach provides a pragmatic analysis of exhaustivity in clefts and definite pseudoclefts which is based on the assumption that both sentence types are anaphoric and introduce an existence presupposition. The proposed analysis generates some interesting issues to be investigated in future research, such as the role of (implicit or explicit) questions in the (non-)exhaustivity of clefts, as well as possible cross-linguistic differences given the differing discourse-semantics for cleft constructions (e.g., German vs. French). The results reported here and the proposal to account for them provide a stepping stone for more theoretical, experimental, and cross-linguistic work in order to get a more fully detailed compositional analysis of the exhaustivity inference in clefts (and definite pseudoclefts).

Chapter 2

On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts*

Abstract One outstanding issue in the analysis of the meaning of clefts concerns the source of the exhaustive inference they convey. Conventionally-coded semantic accounts predict that this inference is robust and will arise regardless of contextual variation while allowing for cross-linguistic variation. On the contrary, non-conventionally-coded pragmatic accounts predict exhaustivity to be more variable within a language, including cases where it can be cancelled, although (potentially) the inference will be more stable across languages. This article presents an original empirical perspective on the debate by looking both at the interpretative and the processing properties of English compared to French clefts. The combination of offline and online measures reported here show crucial and surprising differences within and across the two languages, findings which are unexpected under all current theories of clefts' meaning. We discuss a preliminary sketch for an analysis, which proposes that the differences between French and English are due to the way the existential presupposition derived from the cleft structure interacts with context (cf. Pollard & Yasavul 2016, De Veaugh-Geiss et al. 2018).

* The authors wish to thank Thomas Farmer for his help and advice with analyzing data in earlier versions of the experiments, as well as the XPrag.de program funded by the DFG SPP 172. This paper has greatly benefitted from discussions with many people, including David Beaver, Judith Degen, Mira Grubic, Edgar Onea, Clare Patterson, Agata Renans, Judith Tonhauser, Leah Velleman, Marta Wierzba, and Malte Zimmermann, as well as the members of the audience at the Texas Linguistics Conference 2014, Cognitive Science Conference 2015, and XPrag 2015. All errors are our own.

DOI: https://dx.doi.org/10.1016/j.pragma.2018.09.009

Keywords: French; English; clefts; exhaustive inference; existence presupposition; response times

2.1 Introduction

In English, in addition to asserting the proposition (2a) and carrying an existential presupposition (2b), the focus-background *it*-cleft in (1) also triggers an exhaustive inference such that the pivot is interpreted as if under the scope of an exclusive particle (2c).

(1) It is baby who is shaking a rattle.

(2) a. A baby is shaking a rattle. (prejacent proposition)
b. Someone is shaking a rattle. (existential presupposition)
c. Only a baby is shaking a rattle. (exhaustive inference)

One outstanding problem in the literature on the meaning of clefts concerns the source of this exhaustivity. Opinions differ mainly along a semantic-pragmatic divide, boiling down to whether the inference is encoded as part of the conventional meaning of clefts (Büring & Križ 2013, Velleman et al. 2012) or whether it is derived from pragmatic reasoning on the context (Horn 1981). Cross-linguistically, similar structures (at least in surface) are also acknowledged to convey exhaustivity. Specific cases of this are the Hungarian pre-verbal focus position (É. Kiss 1998), the German *es*-cleft (Drenhaus, Zimmermann & Vasishth 2011), and of core interest for this paper, the French *c'est*-cleft in (3) (Lambrecht 1994).

(3) C'est un bébé qui agite un hochet.
it-is a baby who shakes a rattle
'*It's a baby who is shaking a rattle.*'

One question is whether the exhaustive effects in these different structures are expressed with the same strength and systematicity. From a theoretical perspective, the semantic and pragmatic accounts put forward in the past literature, though mainly developed around English, should in principle be expandable to explain speakers' inferencing behavior with corresponding structures cross-linguistically. Yet to date, there have been few attempts to directly compare the inference across languages, and especially across languages that differ in their use of clefting as a strategy to mark focus (but see Destruel et al. 2015, Skopeteas & Fanselow 2011).

Given this, the main goal of this paper is to provide additional evidence to the debate on modeling exhaustivity by adopting a cross-linguistic perspective. Our general working hypothesis is that speakers of languages with broad uses of clefts will exhibit less robust exhaustive effects, and that differences among speaker's inferential behavior are expected to arise. Two relevant languages to test this hypothesis are English and French. The reason here is that these two languages differ in the options they allow to mark narrow focus (especially on grammatical subjects) and the contexts in which clefts can appear. That is, *it*-clefts are generally marked in English, i.e., preferred in contexts that convey meanings such as contrast (Destruel & Velleman 2014, Destruel, Beaver & Coppock 2017) or correction (Pollard & Yasavul 2016). On the other hand, *c'est*-clefts are more flexible in terms of their function and are used more commonly in French, in which they signal informational and identificational focus, in particular in place of prosodic subject focus (see, among others, Féry 2013, Lambrecht 1994), as well as broad-focus. As a result, our hypothesis predicts that French *c'est*-clefts will exhibit less robust exhaustive effects than English *it*-clefts. We test this prediction by using a sentence-picture verification task that combines offline (truth-value judgments) and online (response time) measures. The current study makes a novel methodological contribution, given that online measures are quite scarce in the literature on the meaning of clefts.

The remainder of the paper is structured as follows. Section 2.2 offers a brief review of the background literature on clefts, in which we further detail the differences between French and English clefts. In this section, we present the most influential theoretical perspectives on the meaning of clefts and the empirical landscape that has ensued from testing the theoretical claims, we review the major accounts on processing of other related inferences, and finally, we make explicit our research questions and hypotheses. We present our experiments and their results in Section 2.3. We provide a general discussion of our results in Section 2.4, and we discuss a way to think about the puzzle they present in Section 2.5. We conclude the paper in Section 2.6.

2.2 Background

2.2.1 Contrasting French and English clefts

There is at least some initial support for the idea that French *c'est*-clefts are similar to English *it*-clefts in meaning. Indeed, prior literature has commonly noted that *c'est*-clefts come with an existential presupposition and convey exhaustive effects (DeCat 2007, Katz 1997, Lambrecht 1994). Despite empirical work on French being

scarce, Destruel (2013) and Destruel et al. (2015) suggest that *c'est*-clefts are indeed somehow exhaustive—though to a lesser extent than exclusives like *seulement* 'only'. Therefore, nothing precludes existing theoretical accounts on English (see Section 2.2.2) to extend to French. But, there are some subtle and crucial differences that set the English and the French clefts apart—thus several reasons that such accounts would not extend to French.

First, French *c'est*-clefts are used more commonly than its English counterpart (Carter-Thomas 2009, Katz Bourns 2014), in particular in comparison to canonical sentence forms (SVO). This is primarily due to constraints on French prosody: whereas English can shift prosodic prominence to match the location of the focus constituent, French is more rigid, placing prosodic stress only at the right edge of an intonation phrase. The *c'est*-cleft, despite adding syntactic complexity, circumvents this prosodic restriction by creating an extra intonation boundary that can align with the focus constituent (Hamlaoui 2009). Consequently, the *c'est*-cleft constitutes the default strategy to signal the simpler focus known as information focus—instantiated in answers to *wh*-questions—especially on grammatical subjects, as in our experimental material (see Section 2.3).[1]

By comparison, the *it*-cleft constitutes a marked structure in English and is typically judged as a 'bad' answer to direct questions. For instance, Destruel & Velleman (2014) find that English speakers are very unlikely to produce an *it*-cleft (versus a canonical SVO sentence) and are also similarly unlikely to rate the cleft as a natural response in contexts where the preceding discourse includes an (overt) *wh*-question such as in (4). Instead, *it*-clefts are shown to be preferred in contexts that license a stronger type of focus known as identificational, contrastive, or corrective focus (É. Kiss 1998, Pollard & Yasavul 2016), as illustrated in the related example (5) from Destruel & Velleman (2014).

(4) A: This bean dip is fantastic. I really want to get the recipe. Who made it?
B: #It's Tim who made it.

(5) A: This bean dip is fantastic. I really want to get the recipe. I can't believe Shannon made it, she's usually not a very good cook.
B: Actually, it's Tim who made it.

Most crucially, the English and French cleft constructions differ in terms of the contexts in which they can felicitously appear. Indeed, *it*-clefts cannot signal *broad* focus. That is, sentences in which no content is given, and in which all information is

1 Lambrecht (1994) argues that canonical sentences with prosodic prominence, while being grammatically well-formed, are pragmatically odd in spoken French in focus-related contexts and occur very rarely. This idea is empirically substantiated; see, among others, Destruel (2013) and Féry (2013), who discuss this focus-marking asymmetry.

new and unknown to the hearer. Moreover, the question corresponding to an English *it*-cleft has to match the focus-background structure of the cleft, thus leading to a (semi-)strict relationship between the cleft and the question it can answer (Abrusán 2016). French *c'est*-clefts, on the contrary, can answer broad-focus questions (Katz 2000, Lambrecht 2001), shown in (6) from Clech-Darbon, Rebuschi & Rialland (1999), in which the answer to the question for the cleft of the form '*It is X who Z*' is not congruent to a question derived from the cleft relative, i.e., '*Who Zed?*'—or a sub-question of this question—but rather the much broader question '*What happened?*'

(6) Q: Qu'est-ce qu'il s'est passé?
what is it that REFL.3.SG is happened
'*What happened?*'
A: C'est le petit qui est tombé dans l'escalier.
it is the small-one who is fallen in the stairs
'*The little one fell down the stairs.*'

In sum, French clefts are used both more commonly, in particular in place of canonical SVO sentence forms, and more broadly (i.e., in more focus-contexts) than their English counterparts, but they are nonetheless noted to convey an exhaustive inference. Thus, the existing analyses on English should in principle apply to analyzing exhaustivity in *c'est*-clefts as well.

2.2.2 Past theoretical accounts on cleft exhaustivity

Although many constructions across languages can convey exhaustivity, much of the past theoretical literature has been developed around introspective judgments on the English *it*-cleft (see among others, Atlas & Levinson 1981, Büring & Križ 2013, Horn 1981, Velleman et al. 2012), empirical evidence has only arisen in recent years. Two opposing approaches have been proposed. Either exhaustivity in clefts is treated as a conventionally-coded semantic phenomenon (Atlas & Levinson 1981, Büring & Križ 2013, Velleman et al. 2012), or as an instance of pragmatic enrichment (Horn 1981, 2014, De Veaugh-Geiss et al. 2015, Pollard & Yasavul 2016).

Most semantic analyses of cleft exhaustivity argue that exhaustivity is in some way presupposed (although see Atlas & Levinson 1981, in which exhaustivity is taken to be part of the asserted truth-conditions of the cleft sentence). The work of Percus (1997) and Hedberg (2013) hold that cleft sentences contain a covert determiner element, or some more complex compositional derivation, that makes them semantically equivalent to definite descriptions, e.g., *The one who is shaking a rattle is a baby*. These are assumed to be semantically exhaustive. In a similar vein, Büring & Križ (2013) offer an analysis in terms of a homogeneity presupposition.

According to this account, the cleft denotation must not be part of a larger sum individual satisfying the backgrounded predicate. Crucially though, all semantic accounts contend that exhaustivity effects in clefts are directly derived from the syntactic configuration. Put differently, they are part of the conventional meaning of the cleft itself. Such accounts make several clear empirical predictions. First, the interpretative effects of clefts within a language should be robust and systematic, i.e., they will arise whenever the syntactic structure is encountered in discourse, regardless of context. Second, they cannot be (easily) cancelled. Finally, of core interest for this paper, is the prediction that variation may arise between the two languages tested. Indeed, since felicity-constraints may differ depending on language-specific semantic coding in the cleft structures themselves, the degree of exhaustivity attributed to clefts may differ across speakers of these different languages.

In opposition to this view, pragmatic accounts of exhaustivity, Horn (1981, 2014) take the inference to be a generalized conversational implicature (GCI). According to Grice (1975) and, later on Levinson (2000), GCIs are taken to arise as a matter of default, but, because they are not part of the meaning explicitly endorsed by the speaker (i.e., the asserted meaning), they can be cancelled if not supported by the context. De Veaugh-Geiss et al. (2015) also present a pragmatic analysis of exhaustivity in clefts, in which the inference is a focus-triggered scalar implicature. They argue that clefts are a structural device for marking focus unambiguously and they give rise to stronger exhaustivity effects than their canonical counterparts.

Recently, Pollard & Yasavul (2016), present a dynamic account of exhaustivity illustrated in Section 2.5. They argue that exhaustivity is not coded in the cleft *per se*, but rather is the result of the interaction of the existence presupposition of clefts with the meaning of *wh*-questions (Hamblin 1971). In this account, the existence presupposition of clefts are anaphoric (e.g., Delin 1992), and the exhaustive/non-exhaustive interpretation comes about in how the antecedent discourse referent is resolved (Pollard & Yasavul 2016, De Veaugh-Geiss et al. 2018). In the non-exhaustive case, clefts pick up some (non-maximal) discourse referent to designate further. This can be illustrated in the case of correction, e.g., when revising misinformation about a referent in the discourse, as in (7), adapted from Pollard & Yasavul (2016).[2]

(7) A: Did you hear, Bob got an NSF grant!
B: Well, actually, it was John. And Mike got one, too!

2 One could claim that the acceptability of the second clause in B's response is an example of domain widening; however, in the same discourse with the exclusive *only* instead of the cleft, the continuation becomes unacceptable.

When the cleft answers a *wh*-question, exhaustivity arises: (i) the *wh*-question introduces a maximal discourse referent, and (ii) the cleft existential has this discourse referent as its antecedent.

Regardless of how exhaustivity is exactly derived, all pragmatic accounts make the same clear predictions. First, exhaustivity in clefts is subject to defeasibility, and the content of these inferences can be reasserted by the speaker without giving a feeling of redundancy. Furthermore, assuming that the mechanisms that derive inferences such as implicatures are somehow universal or generalizable, one should expect little to no variation across languages. In other words, all speakers should derive exhaustivity in clefts with the same strength, thus exhibiting similar inferencing behavior.

In the next subsection, we explore how the predictions from the theoretical approaches have fared in light of recent empirical evidence.

2.2.3 The empirical landscape

In recent years, an emerging body of experimental work has posed several challenges to strict semantic accounts of exhaustivity in clefts. Much of this work has relied on exhaustivity violations to test for the nature of the inference, that is, by comparing the behavior of the exclusive particle *only* to clefts, and other strong focus positions. The linking hypothesis behind these studies is that, if exhaustivity is violated but can nonetheless be cancelled in the case of unembedded clefts, then it must not be semantically encoded. Indeed, findings suggest that the exhaustive effects observed with clefts are more easily cancellable than exclusives (see, e.g., Destruel et al. 2015, Onea & Beaver 2009, Xue & Onea 2011), although they might be less easily cancellable than for the corresponding SVO structures (for German: De Veaugh-Geiss et al. 2017, 2018; for English and German: Zimmermann et al. to appear). Moreover, results show that the cancellation of exhaustivity with clefts is at least marginally acceptable in felicity judgment tasks (Byram-Washburn, Kaiser & Zubizarreta 2013, Saur 2013, De Veaugh-Geiss et al. 2015).

A further challenge for semantic analyses is that exhaustivity in clefts and other fronting strategies has been found to interact with contextual factors (Gerőcs, Babarczy & Surányi 2014, Skopeteas & Fanselow 2011) as well as vary cross-linguistically (Onea & Beaver 2009, Destruel 2013, Hole & Zimmermann 2013, Skopeteas & Fanselow 2011). For instance, Onea & Beaver (2009) compared Hungarian preverbal focus and German prosodic focus (i.e., sentences bearing an A-accent). They found that the former—the structural focus position—was associated with stronger exhaustive effects. In a similar vein, a recent study by Zimmermann et al. (to appear) reports on several verification/falsification tasks in Hungarian, German, and English. They found distributional differences in the exhaustive inference between the prever-

bal focus position in Hungarian and clefts in German and English. A further study by Skopeteas & Fanselow (2011) compared languages that all have a left peripheral focus position (i.e., Spanish, Greek, and German), and showed that exhaustivity is significantly weakened when the fronted focus element is *unexpected* relative to more likely alternatives. The same study also demonstrated that the Hungarian preverbal focus position displays more robust effects than the corresponding position in Spanish, German, and Greek. Also, Gerőcs, Babarczy & Surányi (2014) found that the exhaustive effects associated with the Hungarian preverbal focus position was weaker when no explicit question was present, and the amount of time participants had to respond decreased.[3]

Last, a recent study by Tieu & Križ (2017) on the L1 acquisition of exhaustivity indeed hints at differences between English and French. Existing data on the acquisition of English *it*-clefts suggests that children start out by interpreting clefts non-exhaustively and have partly acquired exhaustivity around the age of 4–5 years old (Heizmann 2007, 2012).[4] In Tieu & Križ's (2017) truth-value judgment task, children looked at pictures containing three familiar objects (created in an exhaustive and a non-exhaustive condition) while a puppet described them in a video using a cleft sentence, an exclusive sentence, or a sentence with a definite description. Children were then asked to judge whether the sentence uttered by the puppet accurately described the picture or not. However, similar to English, French-speaking children started out by interpreting clefts non-exhaustively, they were found to continue interpreting clefts non-exhaustively at 6 years old (i.e., comparatively later than English-speaking children in Heizmann's studies).

In sum, while the theoretical literature trends toward supporting semantic analyses of exhaustivity in clefts, the empirical literature has largely been compatible with non-truth-conditional accounts. Most offline measures would however benefit by corresponding online data, providing insight into underlying cognitive processes. Indeed, the linking hypothesis here is that the cognitive operations underlying the cancellation of the exhaustive inference are costly. In other words, if cleft exhaus-

3 Importantly, we must note that pragmatic accounts of cleft exhaustivity are not immune to problems either. One limitation is that the operation of cancellation is not always straightforward, and there is in fact much variation in judgments on the perceived (un)acceptability of added information with clefts. Another limitation is that the validity of the link between the robustness of the exhaustive inference and its semantic-pragmatic source merits being called into question. For instance, Spector (2014) and Bade (2015) present cases where implicatures appear to be obligatory, thus breaking at least one direction of the equivalence between cancellability and implicature.

4 We thank an anonymous reviewer for pointing out that there also exists experimental literature on the L1 acquisition of the exhaustive interpretation in Hungarian focus, which converges on the finding that children before the age of 6 do not associate an exhaustive interpretation with the Hungarian pre-verbal focus position (see, among others, Pintér 2015, 2016).

tivity is activated by default, but cancellation occurs in a second costly step, longer processing times should arise in contexts where exhaustivity is violated.

Because none of the current theories on clefts' meaning make clear predictions regarding the processing of the exhaustive inference, the next subsection turns to examining the psycholinguistic literature on the time-course of two crucial non-truth-conditional components of meaning, implicatures and presuppositions.

2.2.4 On the processing on related inferences

One of the main concerns in the literature on processing inferences is to understand how these non-truth-conditional components are computed with respect to the truth-conditional content in a given expression. The most extensive investigations conducted involve scalar implicatures, notably, the implicature from quantifier 'some', whereby a sentence such as (8) literally encodes (8a) but is also taken to imply (8b) (see Bott & Noveck 2004, Breheny, Katsos & Williams 2006, Huang & Snedeker 2009, Grodner et al. 2010, Nieuwland, Ditman & Kuperberg 2010, Tomlinson, Bailey & Bott 2013).

(8) Some students attended the conference.

a. *At least one* student attended the conference.
(lower-bound semantic interpretation)

b. *Some but not all* students attended the conference.
(upper-bound pragmatic interpretation)

Historically, two opposing views have been proposed. On the one hand, the Default hypothesis (Chierchia 2004, Levinson 2000), argues that the (generalized conversational) implicature in (8b) arises automatically and effortlessly within the interpretation of the sentence. Likewise, this should happen independent of context. This predicts that the upper-bound pragmatic interpretation should be less resource-intensive and therefore faster than the lower-bound semantic interpretation. On the other hand, the Literal-First hypothesis (Huang & Snedeker 2009) posits that the semantic interpretation in (8b), compatible with *all*, is computed rapidly as a by-product of basic sentence processing, which is then negated to arrive at the enriched meaning. Therefore, this account predicts that scalar implicatures require extra time and resources, thus making the opposite prediction from the Default hypothesis.

Empirically, though, a few studies suggest that scalar implicatures are accessed rapidly and produce no obvious processing cost (see Grodner et al. 2010, Breheny, Ferguson & Katsos 2013). The majority of the studies have lent support to the Literal-First account, providing evidence that the derivation of scalar implicatures do incur processing costs. This has been replicated across methodologies, e.g., truth-

value judgments (Bott & Noveck 2004), self-paced reading (Breheny, Katsos & Williams 2006), and eye-tracking (Huang & Snedeker 2009, Storto & Tanenhaus 2005). As summed up by Huang & Snedeker (2009: p. 408), "even the most robust pragmatic inferences take additional time to compute."

Research on the processing of presuppositions, in contrast, is relatively nascent and suggests that the presupposed content of an utterance is available and integrated very rapidly; in some cases, on a par with asserted content (Schwarz 2014). This has been found in a variety of experimental measures, including self-paced reading tasks (Altmann & Steedman 1988), ERP studies (van Berkum et al. 2003, Burkhardt 2006), and eye-tracking studies (Schwarz 2014, 2015).[5] One particular relevant point for the experiment reported here is that very few studies with online measures of presupposition violations can be found in the literature. One example is Tiemann et al. (2011), who found faster reading times in regions following a presupposition trigger for contexts which explicitly contradicted the presupposition in comparison to both neutral and supporting contexts. This finding suggests that participants had processed the violation, but had given up on parsing the remainder of the sentence.

2.2.5 Research questions and hypotheses

The experiments reported hereafter are motivated by three observations: (i) the gap between the theoretical and the empirical literature on clefts; (ii) the limited amount of work on the meaning of clefts with a direct cross-linguistic comparison; and, to date (iii) the little systematic empirical evidence with an online component. Consequently, we aim to address the questions below:

I. Are there differences in terms of the strength and systematicity of the exhaustivity interpretation for clefts within and across English and French, specifically in comparison to exclusives and intonational focus constructions?

II. Are the underlying processing costs involved in the derivation of exhaustivity the same across sentences and languages?

As such, our data can provide a baseline for participants' relative preference of accepting/rejecting clefts across languages when exhaustivity is violated in context, and a baseline to evaluate whether the preference is reflected in online processing (by analyzing participants' response times).

5 However, similar to the recent experimental studies on implicature processing suggesting that contextual cues play a significant role (Degen 2015, Degen & Tanenhaus 2016), it is worth noting that Kim (2007) finds that attention to presuppositions partly depends on context, and argues assertions and presuppositions are processed differently.

As discussed in Section 2.2.3, a general finding in offline behavioral studies is that, in contexts violating exhaustivity, clefts elicit different response patterns relative not only to the asserted exhaustivity of exclusives, but also to the pragmatic exhaustivity of prosodic focus constructions, at least in English. Thus, we expect such a difference both to be found in offline measures (truth-value judgments) and to be reflected in online measures (response times) as well. Furthermore, in light of differences in the felicitous discourse contexts for French and English clefts, we expect that French *c'est*-clefts will exhibit less robust exhaustivity effects in violation contexts. Correspondingly, we also expect clefts in French to not exhibit the same degree of processing costs in these same contexts as clefts in English. Finally, we expect that, if the exhaustive inference is automatic or default in a given language, then processing times should favor such parses, which will lead to differences in response times when participants reject ('false') or accept ('true') the cleft as appropriate in contexts violating exhaustivity.

Overall, the results will bear on the debate on the nature of exhaustivity in clefts by providing data from a novel test that directly compares speakers' inferential behavior in two languages. Specifically, if differences are indeed empirically substantiated, this can tell us whether French clefts should still be treated on par with English clefts, or whether we should posit a different meaning entirely. In the end, we sketch out a unified account in Section 2.5.

2.3 The experiments

The experiment—a sentence-picture verification task—was conducted in English and French. Because the design, procedure, and material are similar in all versions of the experiment, we present the common elements of the methodology in 2.3.1. We then discuss the results per language per language, in 2.3.2.1 for English, and in 2.3.2.2 for French.

2.3.1 Methods

2.3.1.1 Participants

For English, a total of 64 undergraduates from a Midwestern university (age range 18–21 years old), all native monolingual speakers, were given extra credit for their participation. For French, a total of 64 native monolingual speakers (all from Southwestern France) were recruited and given monetary compensation for their participation. Of these participants, 89% were undergraduate students, 9% graduate students, and 2% young professionals working at a university. All participants were

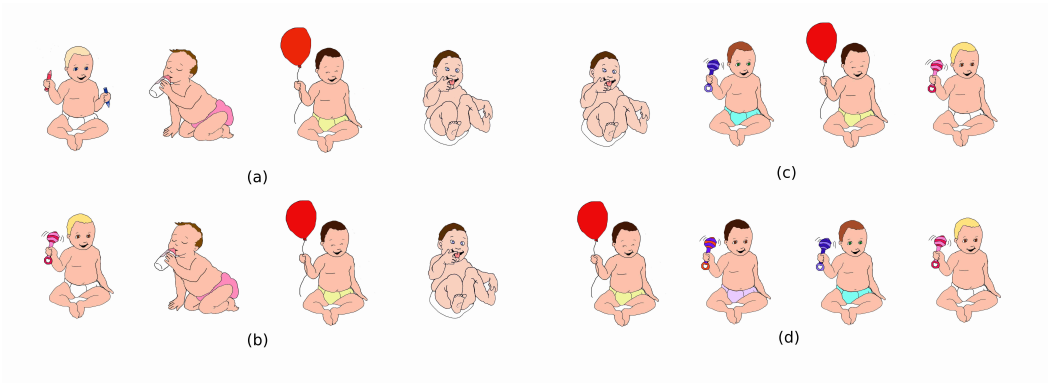


Figure 2.1 *Sample pictorial stimulus.*

naïve as to the purpose of the experiment and had normal or corrected-to-normal hearing and vision.

2.3.1.2 Materials and design

We manipulated two within-subject factors. The first was the *Picture* type seen by participants, which had four levels/conditions (see Fig. 2.1), and for which we created forty different versions per condition.

In the WRONG condition in (a), none of the four characters on the picture has the property described by the predicate. In the –VIOLATION condition in (b), one character has the property asserted in the sentence, supporting the inference '*no one other than X has property Z*'. Finally, in the main condition of interest, the +VIOLATION conditions in (c) and (d), at least one alternative character is also performing the described action, such that exhaustivity is violated.[6] Crucially, all pictures consisted of four characters of roughly the same size, color, and shape (unless otherwise required by the descriptive adjective in the sentence), and in order to avoid recognition effects, the location of the target character was counterbalanced across the four positions in the picture.

The second factor manipulated was the *Sentence* form, for which we created forty lexicalizations for each of the three conditions (CLEFT, EXCLUSIVE, and SVO sentences with prosodic focus). An illustrative test item is given in (9) (see also Appendix B.1 and B.2).

6 A post-hoc analysis revealed no significant difference in truth-value judgments or response times between the 1-actor pictures (Fig. 2.1c) or 2-actor pictures (Fig. 2.1d), either in English or in French, so we collapsed responses together and report on the aggregate of responses for the +VIOLATION condition in Section 2.3.2.

(9) a. It's a [blond]$_F$ baby who is shaking a rattle.
C'est un bébé [blond]$_F$ qui secoue un hochet.
b. Only a [blond]$_F$ baby is shaking a rattle.
Seul un bébé [blond]$_F$ secoue un hochet.
c. A [blond]$_F$ baby is shaking a rattle.
Un bébé [blond]$_F$ secoue un hochet.

All sentences followed the same basic pattern, including the indefinite article *un(e)* 'a', a descriptive adjective, a [+human] subject noun, a transitive verb, and a [–animate] object.[7] Importantly, only a portion of the NP was focused, the adjective (signaled via a pitch accent), which may help in determining the exhaustive effects that arise, or in other words, in identifying the set of alternatives relevant for the interpretation of the sentence.[8] For instance, given the experimental item in (9a), it would be wrong to argue that the open proposition exhaustified is '*X is shaking a rattle*'; instead, the exhaustive effects are sensitive to the specific focus domain. In order to ensure that the desired intonational pattern was achieved, i.e., with focus realized on the adjective, the adult native speakers recording the experimental items were prompted by the experimenter with a *wh*-question of the form '*Which X did Z?*' in the respective language.

In addition to the 40 experimental items, we created 40 fillers (consisting of other non-canonical structures such as existentials, definite descriptions, and passives, all with focus on the adjective), and randomized the total into eight experimental lists via a 2 x 3 Latin square design.

For each language, two versions of the experiment were created. In version 1, the target sentences tested were exclusives and clefts; in version 2, clefts and canonical sentences with prosodic focus. In all other respects, the two versions of the experiments were identical; crucially, the experimental stimuli for the cleft condition in both version 1 and version 2 were exactly the same. In total, 32 participants completed each version of the experiment in English and French (which we will refer to as E1, E2, F1, F2). Thus, overall each participant in each version of the

7 Because in French adjectives can be either pre- or post-nominal, and to control for any effects of word order, the French stimuli included an equal number of items with pre- and post-nominal adjectives.

8 It is important to note that because French is known to less freely resort to prosody to signal focus than English, the prosodic disambiguation of the QUD that comes along with the presence of a pitch accent for English is not as reliable for French speakers; i.e., although prosody can be used to mark informational focus (see Beyssade et al. 2015, Delais-Roussarie et al. 2002, Jun & Fougeron 2000, Welby 2006), it is not clear that French speakers use this as a cue in comprehension. For instance, in an experiment investigating the prosody of DPs, Hamlaoui, Coridun & Féry (2012) find that French speakers do realize different focus structures in a DP in the same way (e.g., [moineau]$_F$ marron vs. moineau [marron]$_F$), yet a follow-up perception test by the same authors revealed that speakers could not accurately distinguish the intended focus structures on the basis of prosody alone.

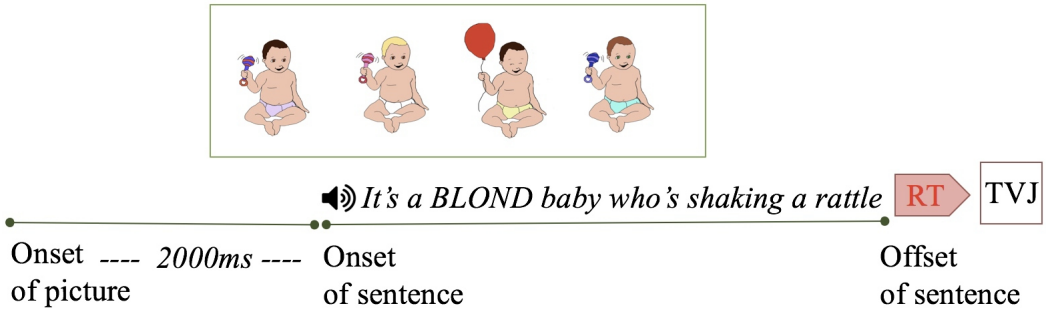


Figure 2.2 *Procedural steps in the experiment.*

experiment judged a total of 80 sentences, only seeing each target item in one of the conditions.

2.3.1.3 Procedure

Participants completed the task, which lasted approximately 30–40 minutes, in a laboratory using the software SuperLab Pro 4.5 (Cedrus Corporation, USA). A training session consisting of two practice trials introduced participants to the task before the experiment began. On every trial, participants looked at a single centrally-located picture displayed on the computer screen for 2000 ms, after which they heard the stimulus in a set of headphones at a self-adjusted volume. The picture remained on the screen as the sentence finished playing. Participants were asked to judge as fast as possible whether the sentence appropriately described the picture by pressing a True or False button (counterbalanced) on a USB Response Pad (RB-530). Between each trial, a white screen appeared for 2500 ms. We collected truth-value judgments (TVJ) and recorded response times (RT). Fig. 2.2 illustrates how the procedure unfolded and what the RTs analyzed correspond to.

2.3.1.4 Data preparation & analysis

Since sentence duration varied across sentence conditions and across language (the English and the French experimental items differed in length), we analyzed RTs from the offset of the sentence, and thus removed from the final analysis all responses made while the sentence was still playing. Furthermore, because participants were permitted to execute their response at any point after the sentence started playing, they could have in principle responded immediately after the onset of the adjective (e.g., *blond*) without considering the remaining portions of the sentence. As such, analyzing RTs from the offset of the sentence also allowed us to exclude results

from participants who may have been guessing or anticipating heuristically what the end of the sentence might be. We should note a potential issue with the French version of the experiment, related to the form of the indefinite article. In French, the indefinite article *un*/*une* has the same form as the numeral *one*. If we find that participants interpret the stimuli as strictly exhaustive, it will be difficult to decide whether they did so because of the meaning they attribute to the sentence structure itself, or because of the fact that they interpreted the sentence as meaning '*exactly one X*' due to the indefinite. To cope with this possibility, we asked participants whether they had interpreted the indefinite as the numeral *one* in a short debriefing session following the experiment. We found that two participants, who completed the second version French experiment (F2), did so, thus consistently judging both types of sentences as 'false' in the +VIOLATION picture condition. We decided to exclude this data from the analysis, so the results for F2 are based on 30 participants instead of 32.

All RT data were log-transformed, and for the *Sentence* form and *Truth-Value Judgment* predictors sum contrasts encoded as numeric covariates were used (*Sentence*: CLEFT: 1, EXCLUSIVE/SVO: –1; *TVJ*: TRUE: 1, FALSE: –1). Note that, although we measured response times for all picture conditions, since we have no particular predictions for the RTs of the other picture types, we only report the results for the +VIOLATION picture, the main condition of interest. We report parsimonious mixed models following the recommendations made in Bates, Kliegl, et al. (2015), including random intercepts and random slopes for participants and items supported by the data, identified utilizing the *rePCA* function in the *RePsychLing* library[9] (MIT, v.0.0.4). For the generalized linear mixed-effects models for the binary TVJ data, we report on estimates, standard errors, z-values, and p- values; and for the linear mixed-effects models for the RT data we report on estimates, standard errors, and t-values, with any t-value exceeding 1.96 considered statistically significant with $p < 0.05$. Analyses were implemented using the *lme4* library in the R environment (GPL-2|GPL-3, v.3.3.3) (R Core Team 2017).

2.3.2 Results

2.3.2.1 English results

Table 2.1 illustrates the TVJ results in percentage for the 'true' judgments, for both versions of the experiment E1 and E2.

The high rate of accuracy for response in both the WRONG and the –VIOLATION control picture condition indicates that participants were engaged in the task and not responding at chance. In the condition of interest (+VIOLATION), 'true' judgments

9 Available at https://github.com/dmbates/RePsychLing.

		+VIOLATION % 'true' TVJ	−VIOLATION % 'true' TVJ	WRONG % 'true' TVJ
E1	*it*-CLEFT	53%	96%	0%
	EXCLUSIVE	1%	99%	1%
E2	*it*-CLEFT	65%	97%	3%
	SVO	87%	97%	1%

Table 2.1 *Percentage of 'true' truth-value judgments for E1 and E2.*

for *it*-clefts were almost evenly divided in E1 (53%) but more biased toward 'true' judgments in E2 (65%), perhaps due to the local effect of the exclusive condition making exhaustivity more salient in version 1 of the experiment. In line with previous experimental studies which found *it*-clefts to show weaker exhaustivity than exclusives yet stronger exhaustive effects than their canonical counterparts, in the +VIOLATION condition participants were significantly more likely to choose 'false' for EXCLUSIVES (E1 $\hat{\beta}$ = 3.4049, SE = 0.5469, z = 6.226, p = 4.79e–10) and 'true' for SVO sentences (E2 $\hat{\beta}$ = –0.9836, SE = 0.3011, z = 3.266, p = 0.00109) when compared to CLEFTS.

Turning to response times, let us first recall that RTs were analyzed from the offset of the sentence until the time when participants pressed the T/F button to indicate their judgment. The two left graphs in Fig. 2.3 show RTs collapsed for all truth-value judgments (i.e., aggregating 'true'/'false' responses) in the +VIOLATION condition. When exhaustivity did not hold, English participants showed a significant delay when making judgments for CLEFTS in comparison to both EXCLUSIVES (E1 $\hat{\beta}$ = 0.15818, SE = 0.07672, t = 2.06) and SVO sentences (E2 $\hat{\beta}$ = 0.10740, SE = 0.05200, t = 2.07). Now zooming in on the cleft condition, we find that participants took significantly more time to give a 'true' judgment than a 'false' judgment in both English experiments, seen in the two right graphs in Fig. 2.3. A mixed-effects logistic regression model predicting log RTs from the truth-value judgment revealed a significant difference in both experiments (E1 $\hat{\beta}$ = 0.2871, SE = 0.0664, t = 4.32; E2 $\hat{\beta}$ = 0.26090, SE = 0.07200, t = 3.62).

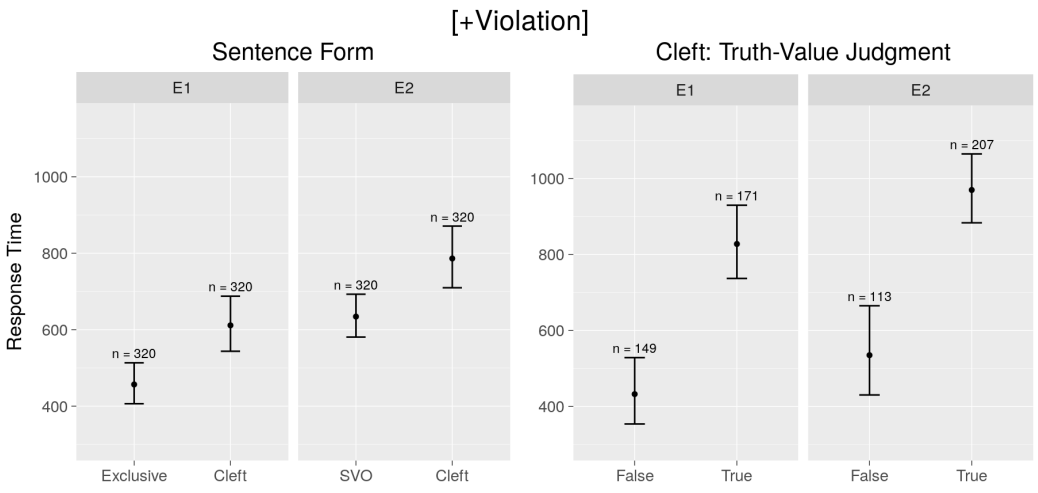


Figure 2.3 *Response Times with 95% confidence intervals (back-transformed means in ms) in* +VIOLATION *condition for English E1 and E2, conditional on Sentence form for all 'true'/'false' responses (left) as well as per Truth-Value Judgment made in the* CLEFT *condition only (right). Total number of judgments appear above the confidence intervals.*

2.3.2.2 French results

Table 2.2 illustrates the TVJ results in percentage for the 'true' judgments, for both versions of the experiment F1 and F2.

		+VIOLATION % 'true' TVJ	−VIOLATION % 'true' TVJ	WRONG % 'true' TVJ
F1	*c'est*-CLEFT	74%	99%	0%
	EXCLUSIVE	2%	97%	1%
F2	*c'est*-CLEFT	78%	97%	2%
	SVO	90%	99%	3%

Table 2.2 *Percentage of 'true' truth-value judgments for F1 and F2.*

We observe that both clefts and canonicals are widely accepted in the +VIOLATION condition, suggesting that participants did not interpret the French indefinite article *un* as the numeral 'one', as discussed in Section 2.3.1.4. As in English, when exhaustivity was violated participants were significantly less likely to select 'true' for EXCLUSIVES than for CLEFTS (F1 $\hat{\beta}$ = 4.6181, SE = 0.7786, z = 5.932, p =

3.00e–09); by contrast, in F2 when comparing SVO and CLEFTS no significant difference was found (F2 $\hat{\beta}$ = –0.5599, SE = 0.7153, z = –0.783, p = 0.434). The descriptively higher number of 'true' TVJs (F1: 74%; F2: 78%) in addition to the lack of statistical significance found when comparing CLEFTS to SVO is compatible with French *c'est*-clefts being merely weakly exhaustive.

As before, we now examine log RTs in the +VIOLATION condition. The two left graphs in Fig. 2.4 show RTs collapsed across TVJs. We observe that French participants' response times were significantly slower for CLEFTS than for EXCLUSIVES (F1 $\hat{\beta}$ = 0.15120, SE = 0.05246, t = 2.88), but not in the same comparison between CLEFTS and SVO sentences (F2 $\hat{\beta}$ = –0.00278, SE = 0.03751, t = –0.07).[10] Concentrating on the +VIOLATION condition for CLEFTS for both F1 and F2, we examined whether log RTs were affected by the truth-value judgment made. Unlike in English, no significant difference in response times was found between TRUE and FALSE judgments for clefts in either experiment (F1 $\hat{\beta}$ = –0.05878, SE = 0.04333, t = –1.36; F2 $\hat{\beta}$ = –0.04847, SE = 0.04724, t = –1.03), shown in the two right graphs in Fig. 2.4.

2.3.3 Interim summary

Two findings are most relevant to the discussion and proposal. First, the relatively high acceptance rate of clefts in contexts that fail to support exhaustivity in both languages, with the acceptability rate being higher for French than for English. Second, the overall slower processing for English clefts when exhaustivity was violated—in particular depending on the ultimate judgment made, with 'true' responses taking significantly longer than 'false' judgments in English, whereas no discernible difference was found in French. Going back to the research questions in 2.5, taken together the findings suggest that variation does occur between the two languages tested, both offline and in the underlying processing cost involved in the computation of exhaustivity. Overall, English and French appear not to convey exhaustivity in cleft sentences with similar strength and systematicity, with French clefts being associated with a weaker inference.

10 One notable difference that we wish to briefly comment on is the fact that response times for SVO vs. cleft sentences showed a delay in English but not in French. We wish to thank an anonymous reviewer for bringing this to our attention. Although we do not have precise predictions concerning the processing behavior of SVO sentences in either language, the delay in English might be linked to the markedness of the cleft structure compared to plain prosodic focus constructions, which was arguably more salient in this version of the experiment due to the explicit comparison being made between *it*-clefts and SVO sentences. That said, the non-delay in French SVO sentences is also interesting, because, assuming these sentences are indeed marked and infrequent compared to clefts (as argued in Lambrecht 1994, and in direct contrast to English), we might expect a potential delay, but in the reverse direction to English, contrary to fact.

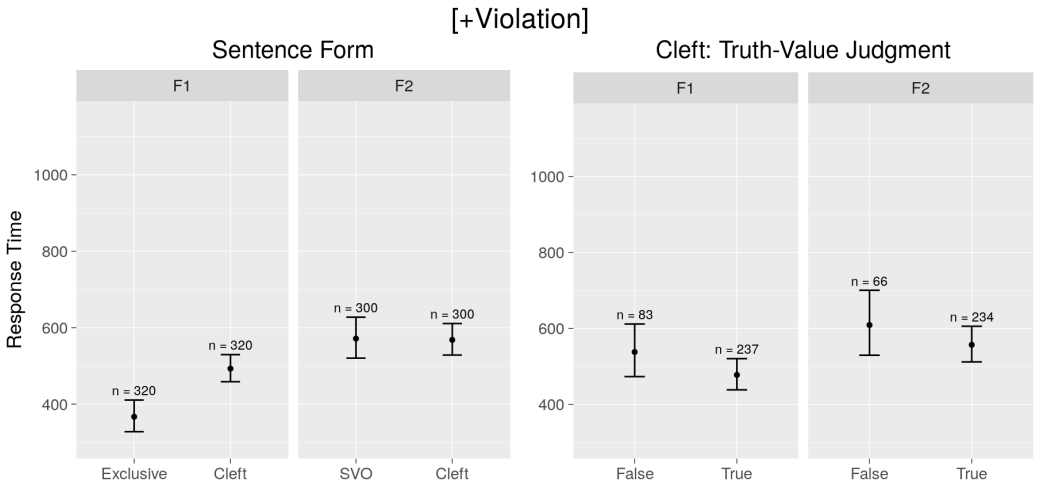


Figure 2.4 *Response Times with 95% confidence intervals (back-transformed means in ms) in* +VIOLATION *condition for French F1 and F2, conditional on Sentence form for all 'true'/'false' responses (left) as well as per Truth-Value Judgment made in the* CLEFT *condition only (right). Total number of judgments appear above the confidence intervals.*

2.4 Discussion

In what follows, we examine our results in more detail, and what we find is that our data presents us with several problems in light of the past theoretical and empirical literature on exhaustivity in clefts. More specifically, in Section 2.4.1 we will compare pragmatic and semantic theoretical approaches to exhaustivity inferences in clefts and conclude that neither can straightforwardly account for the offline results reported above (with the exception of Pollard & Yasavul 2016, which we discuss further in Section 2.5). In Section 2.4.2, we will then look at how the response time measures relate to findings in the processing literature, namely in the processing of implicatures and presuppositions. What we find is that the processing accounts discussed in the literature are incompatible with our results in fundamental ways. In Section 2.4.3 we sum up by returning to the motivation for comparing between French and English, which was at the core of the experiments, before turning to discussing a sketch for a unified proposal.

2.4.1 Theoretical implications

2.4.1.1 Pragmatic theories

At first glance, it seems that pragmatic accounts, as in Horn (1981, 2014), and De Veaugh-Geiss et al. (2015), which treat cleft exhaustivity as a form of pragmatic enrichment, would more straightforwardly be compatible with our offline results for English and French (along with the growing-number of experimental studies showing the relative cancellability of cleft exhaustivity). Indeed, these non-truth-conditional accounts all predict a less stringent version of cleft exhaustivity such that the inference can (more easily) be violated, which is consistent with the relatively high acceptance rates we find for clefts in +VIOLATION contexts.[11]

Upon a first view, pragmatic accounts might appear better suited to explain the online finding for English that exhaustive interpretations elicited quick judgments, perhaps as the default interpretation. Indeed, in Horn's view, exhaustivity is a generalized conversational implicature (i.e., one that arises whenever the cleft is used), which thus arguably arises fairly quickly but can also be subsequently cancelled. Although Horn makes no claim about the underlying processing mechanisms involved in the derivation of exhaustivity with clefts, one tangible hypothesis is that the cancellation step is the one associated with a cost, which is what we find. Compare this, however, to the processing literature which has generally found pragmatic enrichment to be costly. We come back to our findings in terms of processing models in a bit.

2.4.1.2 Semantic theories

One of the core properties of presuppositions is that they are taken for granted: to say that a sentence *S* carries a presupposition *P* roughly means that the use of that sentence is appropriate only if the speaker believes *P* to be part of the accepted common ground for the conversation in which (s)he is involved. Given this property,

11 Similar to our design, the study in De Veaugh-Geiss et al. (2018) had contexts in which exhaustivity was either violated or not. Although we chose not to report on individual variation due to space constraints, it is worth mentioning that our results are in line with theirs: they found participants clustered into one group (consisting of about half the total participants) who treated clefts as exhaustive as exclusives, while another group (consisting of the other half) treated them as non-exhaustive as narrow focus. The authors argued that their results are incompatible with an exhaustivity inference that is conventionally-coded and contextually-entailed, as the semantic accounts of cleft exhaustivity would predict. Indeed, this is where our online data can shed an interesting light on the offline results, especially for English. Our findings suggest that, in English, an exhaustive interpretation might be derived as a default, with its cancellation occurring in a second, costly, step. In principle, this means that although we observe various responder types on the surface (exhaustive vs. non-exhaustive), these speakers might all be similar (exhaustive) to begin with.

there is a strong intuition that presupposition triggers used in contexts that fail to entail *P* will lead to infelicity or falsity. The relatively high acceptance rate of clefts in +VIOLATION contexts clearly represents a challenge for such accounts: when cleft exhaustivity is not supported, the sentence should end up being as unacceptable as other global presupposition failures (see, e.g., Abrusán & Szendrői 2013, Romoli & Schwarz 2015), but we found that this was not (always) the case.

Yet, the observed variation—whereby French speakers appear not to derive an exhaustive interpretation for clefts by default, or that this interpretation is not as strongly activated as with English clefts—is ostensibly less of a challenge for semantic accounts. Indeed, presuppositions do not necessarily have to be homogeneous across languages. For a given presupposition, it is possible for one language to encode it while another does not. This has been argued to be the case, for instance, for the existential presupposition of clefts; while English clefts do encode existence, Straits Salish (Samish) and St'àt'imcets clefts do not.[12] Similarly, Matthewson (2008), comparing English and St'àt'imcets, claims that language variation in the discourse effects of their presuppositions affects the semantics of determiners and third-person pronouns. We also note that strict semantic accounts such as É. Kiss (1998) predict cross-linguistic variation by positing that different languages will specify a positive or negative value for the exhaustive feature encoded in the cleft structure.

In sum, we are left with a puzzling picture, as seen from Table 2.3, showing that our findings cannot be straightforwardly accounted for by any theory of the meaning of clefts in their current forms. In Section 2.5 we will put forward an approach that we believe can best explain our findings, a proposal which—as stated by Pollard & Yasavul (2016), whose analysis we follow—"obviates the need to identify whether this putative implication is a presupposition, a conversational implicature, a conventional implicature, etc". First, however, we will take a closer look at the response time measures as it relates to the literature on implicature/presupposition processing.

	+VIOLATION acceptance	Cross-linguistic variation
Semantic (presupposition)	✗	✓
Pragmatic (implicature)	✓	✗
Our results	✓	✓

Table 2.3 *Predictions of theoretical accounts on clefts' meaning vs. our findings.*

12 See Seth Cable's examples in a handout from his 2008 course "Theoretical Perspectives on Languages of the Pacific Northwest: Proseminar on Semantic Theory." See also the discussion in Tonhauser et al. (2013) about variation in the projective behavior of different presupposition triggers.

2.4.2 Processing literature

2.4.2.1 Implicature processing

Turning to the processing literature, recall that the general trend seems to be that while presuppositions are derived very rapidly, implicatures are generally costly to compute. To the extent that we can take response time measures to represent indexes of processing difficulty, the finding that, in English, participants take less time to provide 'false' judgments in +VIOLATION visual contexts appears to be at odds with the Literal-First hypothesis (i.e., semantic meaning first, pragmatic meaning second) and much of the experimental literature reporting a delay in implicature computation (Bott & Noveck 2004, Breheny, Katsos & Williams 2006, Noveck & Posada 2003). In fact, the English data seems best predicted by the Default hypothesis, in which pragmatic implicatures arise rapidly, automatically, and effortlessly: that is, rejecting the cleft as appropriately describing a non-exhaustive picture is fast because the exhaustive inference is quickly present after the trigger appears, whereas accepting the cleft is costly because of the need to go through (the second processing step of) the annulment of the unsupported inference. Importantly, although some experimental studies do indeed suggest that implicatures are fast to compute (Grodner et al. 2010), they only show no cost of cancellation: at best, the upper-bound pragmatic reading is as fast as the logical, semantic reading, but never faster.

In sum, under traditional approaches to implicature processing the response time measures reported here, in particular for English, are surprising—cleft exhaustivity appears to pattern differently from scalar implicatures. But, we shall make one brief note concerning new directions in the study of implicature processing, which has found that the costs associated with computing implicatures are dependent on several factors (e.g., contextual richness, the availability of alternatives, experimental task, common ground). For instance, recent work by Degen and colleagues (e.g., Degen & Tanenhaus 2016) have proposed a probabilistic account of implicature derivation, re-focusing empirical efforts on examining the role of contextual information sources that contribute to the ultimate interpretation of a speaker's intended meaning (e.g., form of the quantifier, discourse accessibility; see Degen 2015). The common finding in these studies is that scalar implicatures can be modulated, thus arising as a matter of degree depending on support received from various contextual cues. Although our results do not align straightforwardly with traditional accounts of processing, they are conceivably compatible with such accounts: the fact that exhaustivity seems variable and dependent on context is, in principle, expected in these approaches.

2.4.2.2 Presupposition processing

How do our results fit in with research on the processing of presuppositions? We remind the reader that the presupposed content of an utterance is found to be processed and integrated very rapidly in a variety of experimental measures (e.g., Burkhardt 2006, Schwarz 2014, 2015). Since exhaustivity for English *it*-clefts in particular was derived quickly (evidenced by the fast 'false' judgments), an analysis along these lines appears to be a fruitful first step.

As discussed above, however, the offline data poses a puzzle for presuppositional accounts: In Section 2.2.3, we mentioned that the few TVJ and related experimental studies with sentential or contextual presupposition contradiction have found broad rejections or a majority of 'false' judgments (e.g., Abrusán & Szendrői 2013, Romoli & Schwarz 2015). Our visual stimuli similarly contradicted what would be the global contextual entailments predicted by a presuppositional account of exhaustivity, and we expected that if participants are confronted with a violation of this inference they will largely reject the picture. This is not what we found, with the majority of participants instead choosing 'true' in both languages—from 53% (in E1) to 78% (in F2) of the time—in spite of the violation of exhaustivity. Moreover, since presupposition cancellation, local accommodation, or suspension are phenomena which occur in embedding environments, as in the sentences tested in Romoli & Schwarz (2015), standard approaches to presupposition annulment would not predict a presupposition to be cancellable in the types of sentences we tested.

It is important to end our discussion by considering some of the limitations to the conclusions that can be drawn based on the methodology used. All in all, the present methodology does not allow us to decisively differentiate between all analyses on the source of the exhaustive effect in clefts. For instance, we do not know if the cleft sentences were perceived as degraded in the +VIOLATION condition, although past truth-value judgment experiments would predict they are (Saur 2013, De Veaugh-Geiss et al. 2015). Nevertheless, we think—given the quick response times in English for the participants who chose 'false' and delayed response times when judging the sentence as 'true'—the exhaustivity inference was first generated and then for some participants overridden or ignored. So, at this point, additional methodologies are required to make further, stronger claims about the processing of exhaustivity in clefts, such as eye-tracking or mouse-tracking, which would be able to shed more light on the underlying cognitive step(s) participants go through to arrive at a final (truth-value) judgment. What is important for the present paper though is the differences between the offline and online results for English versus French, and not the individual explanations of their integration.

2.4.3 Cross-linguistic perspective

We remind the reader that our initial motivation for testing English compared to French was the relative difference in the use of clefts in the two languages. Should French and English clefts be treated similarly with respect to exhaustivity? The variation observed in our experimental data, with French clefts telling a different story than their English counterpart in both offline and online measures, could lead one to argue that the two are radically different structures, and that an analysis of exhaustivity in English *it*-clefts will have difficulty also accounting for French *c'est*-clefts. In short, instead of treating the two clefts on par with each other, one option would be to posit (completely) different meanings for English and French. However, this strikes us as an easy, and dispreferred, way out. Yet, for a unified approach, whatever analysis one proposes for cleft's exhaustivity in one language must be able to explain the different data in the other. Furthermore, in addition to the differences between English and French in both offline and online measurements, another tension should be explained: while English and French clefts are accepted in non-exhaustive contexts (and this to a larger extent than predicted by theoretical semantic accounts), this acceptability in English—but crucially not French—seems to come at a (processing) cost, the cost of the dismissal of the inference.

2.5 Proposal

Here, we would like to offer a possible solution to the problems that our data represents for current theories of the meaning of clefts. We must acknowledge that we only intend to discuss a sketch of a proposal at this point, and that spelling out more precisely the formal details needed to fully account for (crosslinguistic) variation will be an important research task.

We think that a unified account of exhaustivity in clefts is preferable, and, in a nutshell, we follow an idea that appears in the analyses of Pollard & Yasavul (2016) and De Veaugh-Geiss et al. (2018) (and reminiscent of Horn 1981 in deriving exhaustivity from the existential): the exhaustive inference is derived from an interaction with some other layer of meaning, namely, the existence presupposition of clefts. Crucially, we think the differences in the strength of exhaustivity between French and English could be boiled down to differences in the strictness placed on the requirement for question-answer congruence by both languages. This idea follows the spirit of Abrusán's (2016) account in explaining the soft vs. hard trigger distinction for the existence presupposition in focus and *it*-clefts. Let us now develop our idea slightly.

In our experiment, similar to the design in De Veaugh-Geiss et al. (2018), clefts appeared out-of-the-blue. This absence of context required participants to

accommodate the existence presupposition, which, following Delin (1992), Pollard & Yasavul (2016), and De Veaugh-Geiss et al. (2018), among others, we take to be anaphoric. In this vein, Pollard & Yasavul's dynamic account takes clefts to be devices which pick up an antecedent discourse referent (DR), and whether or not exhaustivity arises depends on how this discourse referent is resolved. That is, Pollard & Yasavul argue that cleft exhaustivity is not coded in the cleft *per se*; rather, exhaustivity arises when clefts are taken to be answering a *wh*-question, and it does not arise otherwise. More specifically, a question accepted in discourse introduces a maximal discourse referent with the property in question. The cleft, when used as an answer, has this maximal discourse referent as its antecedent and is thus interpreted exhaustively. According to this account, however, clefts are not necessarily used as answers to questions, as seen in example in (7), repeated below in (10). In such cases, the cleft picks up some antecedent DR (not necessarily the maximal one) in order to specify it further, and thus an exhaustive reading does not obtain.

(10) A: Did you hear, Bob got an NSF grant!
B: Well, actually, it was John. And Mike got one, too!

As discussed in Section 2.2.1, English clefts have a strict question-answer congruence requirement. As Abrusán (2016) writes, the cleft relative in English "constrains the background question to be the question to which the focused element [in the cleft pivot] is the direct, short answer" (184). This background question is of the form '*Who Z?*' (or a sub-question of this question),[13] which is derivable from the cleft relative '*(It is X) who Z*' itself. That is, taking the denotation of the relative clause to be a lambda-abstract—e.g., for (1), λx.shaking a rattle(x)—it is straightforward to derive the set of propositions in a Hamblin-style question denotation, λp.∃x.[p = λw. shaking a rattle (x) (w)] (Hamblin 1973, Abrusán 2016).

By comparison, French clefts have a less stringent question-answer congruence requirement, and this is exactly the crux of our idea: we argue that for French, the corresponding background question can, but crucially need not be derived from the cleft relative (see, e.g., the all-new focus question in example (6)). How would this idea resolve the differences between French and English? In English, accommodation of the anaphoric existence presupposition out-of-the-blue can go in one of two ways. On the one hand, one can assume that the cleft is an answer to a question, and in this

13 Note that by contrast, a broader range of questions is argued to license plain focus constructions, such as '*Who, if anyone, Z?*' as well as '*Did anyone Z?*', neither question presupposing existence (Abrusán 2016; see also the discussion in Rooth 1996: p. 19). Accordingly, these differences in licensing properties are argued to give focus constructions a weaker existence presupposition compared to both *it*-clefts and preverbal focus in Hungarian, with plain focus in languages such as English being a so-called 'soft' presupposition trigger for existence (Abusch 2002, 2010; see Abrusán 2016 for details).

case the existence presupposition of the cleft will have as its antecedent the maximal discourse referent congruent to the *wh*-question derivable from the cleft relative. On the other hand, the existential is accommodated to a non-maximal discourse referent to specify further. Crucially, we argue that the exhaustive interpretation is the initial default one, since it can be derived (almost) effortlessly from the focus-background structure of the cleft alone given the direct relationship between the cleft relative and the congruent question. A non-exhaustive interpretation will incur further costs, given that the context which would license it requires additional enrichment. Thus, the initial interpretation leads to 'false' judgments, whereas the enriched, i.e., costly, one leads to 'true' judgments. That more participants tended slightly toward the enriched meaning is in line with the claim that clefts most naturally occur in (non-exhaustive) corrective or contrastive contexts (Destruel & Velleman 2014, Destruel et al. 2015). That interpretation, however, comes with a corresponding higher processing cost.

In French, participants may follow a similar process to English, with one major difference: the congruent question is not strictly derivable from the cleft relative, since French clefts can be used to answer a broader range of questions. Thus, when French participants accommodate the existential, they have multiple paths: (i) accommodate a maximal discourse referent answering a question congruent to the backgrounded cleft relative, just as in English; (ii) accommodate a non-maximal discourse referent to specify further, again as in English; or, (iii) accommodate some non-maximal discourse referent relevant in the answer to an all-new focus question such as *What happened?*[14] Note that, of these three options, only one, namely (i), results in an exhaustive interpretation.[15] However, French clefts are more flexible in terms of their function, and unlike in English, the structural cue from which one might derive the background question is ambiguous, since the cleft relative is not strictly congruent to a narrow-focus question. Thus, no obvious and straightforward default strategy to help accommodate the existence presupposition will arise.

14 cf. Onea & Beaver (2016) and De Veaugh-Geiss et al. (2018), in which the indefinite, e.g., *Someone is shaking a rattle* from (1), is argued to give rise to a *potential question* in the sense of Onea (2016), a peculiar non-maximal *wh*-question, e.g., *Who is this person who is shaking a rattle?* Following this line of thought, as one reviewer pointed out, clefts will always answer a question; however, while clefts in English can answer either a maximal *wh*-question or a non-maximal *wh*-question, in French there is still just one maximal question but there is more than one non-maximal question. Along these lines, as the reviewer put it, there may not be a non-question discourse referent a cleft can specify in the end.

15 As one reviewer pointed out, it would be worthwhile looking into how frequency effects might influence our data. Although there have been studies looking at the frequency of the cleft structure across languages (see, e.g., Dufter 2009 for a comparison of Romance languages and German), there are very few directly comparing the distribution of clefts specifically in terms of the meanings in context associated with them (but see Karssenberg & Lahousse 2018 for a pioneering corpus analysis of different types of clefts).

2.6 Conclusion

This paper investigated the relative difference in the strength of exhaustivity associated with two cleft structures: the English *it*-cleft and the French *c'est*-cleft. We employed a picture-sentence verification task for which we analyzed truth-value judgments and response times in different pictorial contexts, and notably one where exhaustivity was violated. The main results were that French and English varied greatly with respect to both factors. French speakers are more readily willing to accept the cleft in contexts violating exhaustivity and doing so without the processing cost that emerged with English speakers. We took these results to suggest that the exhaustive inference is the initial default interpretation in English when no further context is provided, which is not the case in French.

We discussed a sketch of a unified account for cleft exhaustivity based on an idea present in the analysis of Pollard & Yasavul (2016), which proposes that clefts do not encode exhaustivity but that rather, an exhaustive inference may or may not arise depending on how the anaphoric existence presupposition is resolved—either to a maximal discourse referent answering a congruent question (exhaustive interpretation) or to some discourse referent which is then given further specification (non-exhaustive interpretation, e.g., contrastive or corrective). For English speakers, without support of further contextual cues the initial interpretation is the former given the semi-strict relationship between the cleft relative and the congruent question. However, this interpretation can be overridden with contextual enrichment, albeit with a cost. For French, the fact that *c'est*-clefts can be used in broader contexts will make the exhaustive interpretation weaker and not arise as a default interpretation.

Although the work presented here constitutes a modest, yet necessary step towards better understanding the exhaustive inference associated with the English and French clefts, the results open up the possibility for further investigations in two paths. One of the advantages of our proposal is that it predicts that, in languages where clefts are used more broadly in discourse, the exhaustive effects should also be more diluted than as reported in English, Hungarian, and German. In future work on processing, examinations of the time-course for the availability of the exhaustive inference via more robust methods (e.g., eye-tracking or mouse-tracking), and considerations about the influence of other contextual and linguistic factors will also be of great theoretical interest.

Chapter 3

Experimental studies on *it*-clefts and predicate interpretation*

Abstract There is an ongoing discussion in the literature whether the series of sentences '*It's not* α *that did P.* α *and* β *did P.*' is acceptable or not. Whereas the *homogeneity approach* in Büring & Križ 2013, Križ 2016, and Križ 2017 predicts these sentences to be unacceptable, the *alternative-based approach* predicts acceptability depending on the predicate being interpreted distributively or non-distributively (among others, Horn 1981, Velleman et al. 2012, Renans 2016a,b). We report on three experiments testing the predictions of both types of approaches. These studies provide empirical data that not only bears on these approaches, but also allows us to distinguish between different accounts of cleft exhaustivity that might otherwise make the same predictions. The results of the three studies reported here suggest that the acceptability of clefts depends on the interpretation of the predicate, thereby posing a serious challenge to the homogeneity approach, and contributing to the ongoing discussion on the semantics of *it*-clefts.

* Authors listed in reverse alphabetical order. We would like to thank the anonymous reviewers at Semantics & Pragmatics as well as the editor, Kristen Syrett, for the critical and very helpful comments on earlier versions of this paper. We would also like to thank Malte Zimmermann, Jacopo Romoli, Dario Paape, Edgar Onea, as well as the audiences at the "Theoretical and Experimental Approaches to Presuppositions" Workshop at the Università degli studi di Genova, Italy, and the "International Workshop on Non-Prototypical Clefts" at KU Leuven, Belgium. All errors are our own. In order to prevent link rot, all URLs included in this document were archived, when possible, at the Internet Archive's Wayback Machine at https://web.archive.org/. This research was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 317633480—SFB 1287, Project C02, as well as the "Exhaustivity in *it*-Clefts" project in the XPrag.de priority program (SPP 1727).

DOI: https://dx.doi.org/10.3765/sp.12.11

Keywords: *it*-clefts; exhaustivity; homogeneity; distributive, collective, and mixed predicates; distributive vs. non-distributive interpretation; experimental study

3.1 Introduction

It has been observed in the literature that focus-background *it*-clefts of the form '*It's* α *that did P*.' give rise to an exhaustive inference that α in the cleft pivot is the only entity for which the predicate *P* holds (Horn 1981, Percus 1997, Velleman et al. 2012, Büring & Križ 2013, Destruel et al. 2015, Renans 2016a,b, Križ 2017, De Veaugh-Geiss et al. 2018, among many others). For the sake of example, consider the sentence in (1).

(1) It was Kimberly who did the dishes.
⇝ Kimberly and nobody else did the dishes. (exhaustive inference)

There is, however, an ongoing debate as to whether the series of sentences as in (2) is acceptable.

(2) It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.

In particular, the *homogeneity approach* in Büring & Križ 2013, Križ 2016, and Križ 2017 predicts (2) to be invariably unacceptable.[1] By contrast, Velleman et al. (2012) and Renans (2016a,b) observe that the acceptability of (2) will differ depending on the interpretation of the predicate being either distributive or non-distributive; moreover, several other approaches to cleft exhaustivity are compatible with the (un)acceptability of (2) being influenced by the distributive/non-distributive interpretation of the predicate (Horn 1981, Destruel et al. 2015, De Veaugh-Geiss et al. 2018). Despite their differences, we will lump these non-homogeneity accounts into one group which we call, following Križ (2017), the *alternative-based approach*.

As Velleman et al. (2012: p. 455) remark, however: "we must admit that these are subtle intuitions, and further empirical evidence here would be helpful". Therefore, the clarification of the (un)acceptability of (2) is urgent not only for empirical reasons, but more importantly, it provides a new empirical landscape allowing one to distinguish between different approaches to cleft exhaustivity which might otherwise make the same predictions.

1 Büring & Križ (2013: Fn. 1) and Križ (2017: Fn. 1), following Horn (1981, 1985, 1989), acknowledge that it may be possible for (2) to be acceptable if the negation in the first sentence is metalinguistic negation. We come back to this issue in Section 3.4.

The goals of the three studies presented here are twofold: we aim at (i) clarifying the empirical generalizations, and (ii) experimentally testing the theoretical predictions of clefts with distributively and non-distributively interpreted predicates. Results reveal that the distributive vs. non-distributive interpretation of the predicate does indeed influence judgments of acceptability for the series of sentences as in (2): specifically, sentences with distributively interpreted predicates were judged as less acceptable than sentences with non-distributively interpreted predicates, and this effect was robust and replicated across three different experiments. The results reported here are thus consistent with the alternative-based approach to cleft exhaustivity, while posing a direct challenge to the homogeneity approach.

The outline of this paper is as follows: In Section 3.2 we present the theoretical background to both the homogeneity approach and the alternative-based approach, in particular Velleman et al. 2012, along with the predictions of each. In Section 3.3, we present three experiments designed to test these predictions, one with written stimuli and two with auditory stimuli. In Section 3.4, we address the role of negation—in particular, metalinguistic negation—for the reported experiments. Section 3.5 concludes.

3.2 Theoretical background

3.2.1 Predicate types (distributive, collective, and mixed)

We begin by spelling out the theoretical assumptions regarding distributive and other predicate types, namely, collective and mixed predicates. Following Landman 1989, we assume that distributive predicate types predicate of the atomic individuals that constitute the plurality: that is, they have atomic entities in their denotation. For example, upon hearing (3a), one can conclude that Alice laughed, shown in (3b).[2]

(3) *Distributive Predicate*

a. Alice and Bob laughed. $\rightsquigarrow$ *Alice laughed and Bob laughed.*

b. $laughed(a \oplus b) \vDash laughed(a)$ (distributive interpretation)

By contrast, collective predicate types predicate of pluralities only: that is, they only have plural individuals in their denotation. This means that the predicate is not true of each atomic entity. For that reason, given that Carol and Dan gathered, as in (4a), it is not true that Carol gathered, shown in (4b).

(4) *Collective Predicate*

a. Carol and Dan gathered.

2 See Champollion (to appear) for an overview and discussion on distributivity vs. collectivity.

b. $gathered(c \oplus d) \not\models gathered(c)$ (non-distributive interpretation)

Finally, in addition to distributive and collective predicates, there are mixed predicate types such as *do the dishes* in example (5) below. These are referred to as 'mixed' since the interpretation of the predicate depends on the context in which the predicate appears. For example, while in the context of (5a) *do the dishes* is interpreted distributively, in (5b) it is interpreted non-distributively.

(5) a. *Distributive Interpretation*
Helen and Kimberly live together. On Saturday Kimberly did the dishes and yesterday Helen did the dishes.
(i) Kimberly and Helen did the dishes. $\rightsquigarrow$ *Kimberly did the dishes and Helen did the dishes*
(ii) $did\_the\_dishes(a \oplus b) \models did\_the\_dishes(a)$

b. *Non-Distributive Interpretation*
Helen and Kimberly live together. Yesterday Helen cooked, but they did the dishes together.
(i) Kimberly and Helen did the dishes.
(ii) $did\_the\_dishes(a \oplus b) \not\models did\_the\_dishes(a)$

[Experiment 3, Item 12]

There is an ongoing discussion whether such mixed predicates should be analyzed as being underspecified (Kratzer 2008, Schwarzschild 1993, Moltman 1997) or ambiguous (Heim 1994, Moltman 2005, Frazier, Pach & Rayner 1999). In our study, we manipulated the context such that only one interpretation of the predicate—either the distributive or a non-distributive—was pragmatically plausible. Since the interpretation was controlled for by the context, we stay neutral in the debate as to whether mixed predicates are underspecified or ambiguous.

Before we go into the details of our three experiments, let us first discuss two competing approaches to exhaustivity in clefts: the homogeneity approach and the alternative-based approach.

3.2.2 Homogeneity approach

Under the homogeneity approach, *it*-clefts semantically correspond to copular sentences with the cleft predicate turned into a (number-neutral) definite description: they are identity statements between two individuals (Büring & Križ 2013, Križ 2017).[3] The core idea is that the violation of the exhaustivity inference results in

3 While there are in fact differences between the accounts in Büring & Križ 2013 and Križ 2017, they do not influence the predictions for our studies. One difference to note is that under Križ's (2017) formulation, unlike in Büring & Križ's (2013) account, homogeneity in clefts is not a presupposition

the cleft sentence—as well as its negated version—being neither true nor false; that is, there is a truth-value gap. This trivalent property of the predicate is called homogeneity. For example, the sentence in (6) is argued to give rise to the following truth-conditions.

(6) It was Kimberly who did the dishes.

true iff Kimberly is identical to the mereological sum of people who did the dishes
≈ *iff Kimberly and only Kimberly did the dishes*

false iff Kimberly does not overlap with the mereological sum of people who did the dishes
≈ *iff Kimberly did not participate in doing the dishes*

undef otherwise
≈ *iff Kimberly and somebody else did the dishes*

Crucially, negation switches truth and falsity but leaves undefinedness intact.

(7) It wasn't Kimberly who did the dishes.

true iff Kimberly does not overlap with the mereological sum of people who did the dishes
≈ *iff Kimberly did not participate in doing the dishes*

false iff Kimberly is identical to the mereological sum of people who did the dishes
≈ *iff Kimberly and only Kimberly did the dishes*

undef otherwise
≈ *iff Kimberly and somebody else did the dishes*

Now, it is easy to see why the series of sentences '*It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.*' is invariably predicted to be unacceptable under the homogeneity approach. In the situation in which Kimberly and Helen did the dishes, the cleft sentence '*It wasn't Kimberly who did the dishes.*' cannot be true since Kimberly does in fact overlap with the mereological sum of people who did the dishes. It also cannot be false, however, since Kimberly is not identical to the mereological sum of people who did the dishes (i.e., Helen did the dishes as well). Hence, it is undefined. Crucially, since under the homogeneity approach the truth conditions of cleft sentences are defined by overlap with or identity to the

---

and it is not triggered by the definite article but by the predicate. However, as Križ (2017: p. 25) writes: "In terms of actual empirical predictions about clefts in particular, [Büring & Križ's] theory is in some places ill-defined, but once this is remedied in the natural way, the predictions turn out to be entirely identical to ours. The theory we have presented is thus simply to be viewed as an update of [Büring & Križ 2013] [...]".

mereological sum of entities who satisfy the cleft predicate, it makes exactly the same predictions for distributively and non-distributively interpreted predicates.

3.2.3 Alternative-based approach

A contrasting account is the alternative-based approach, which postulates that the meaning of *it*-clefts is fully described by (at least) two meaning components: (i) the asserted canonical meaning of clefts, which is the same as its non-cleft Subject-Verb-Object (SVO) counterpart, and (ii) the non-asserted exhaustive meaning component (e.g., Horn 1981, Destruel et al. 2015, De Veaugh-Geiss et al. 2015, Renans 2016a).[4]

(8) It was Kimberly who did the dishes.
   a. Asserted meaning component:
   Kimberly did the dishes.
   b. Non-asserted meaning component:
   Kimberly and nobody else did the dishes.

The various theories differ in terms of how the non-asserted exhaustive meaning component comes about. For concreteness, we will discuss in detail the account in Velleman et al. 2012, which analyzes clefts as focus-sensitive inquiry-terminating operators indicating a complete answer to the Question Under Discussion (QUD) (Roberts 1996). In particular, clefts assert $\text{MIN}_S(p)(w)$ and presuppose $\text{MAX}_S(p)(w)$, shown in (9).[5]

(9) $\text{CLEFT}_S = \lambda w.\lambda p : \text{MAX}_S(p)(w).\text{MIN}_S(p)(w)$
   a. Asserted (MIN): There is a true answer at least as strong as p.
   b. Presupposed (MAX): No true answer is strictly stronger than p.
   [based on ex. (22) in Velleman et al. 2012]

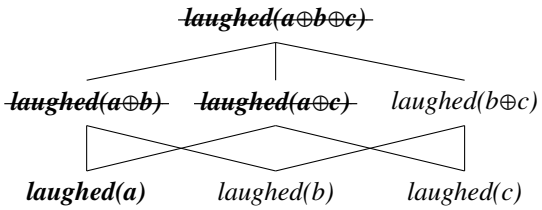
For the sake of illustration, consider both example (10) and the figure in (11), which shows the contextual alternative answers in the entailment scale corresponding to the QUD '*Who laughed?*' (example from Velleman et al. 2012: §4.1).

(10) It was Alice who laughed.

4 We ignore here another meaning component of *it*-clefts, the *existence presupposition*, from which various approaches (e.g., Pollard & Yasavul 2016, Destruel & De Veaugh-Geiss 2018, De Veaugh-Geiss et al. 2018; see also Horn 1981) derive the exhaustive meaning. In direct contrast, Büring & Križ (2013: §6) argue that the existence presupposition is independent of the exhaustivity inference, if it is encoded in *it*-clefts at all.

5 Following Heim & Kratzer's (1998) convention, the presupposed material is between the semicolon and the dot.

(11)



In a domain including three individuals, Alice, Bob, and Carol, MIN asserts that Alice laughed, i.e., that there must be a true answer in the following set, illustrated in (11) by boldface text.

$$\{laughed(a),\ laughed(a\oplus b),\ laughed(a\oplus c),\ laughed(a\oplus b\oplus c)\}$$

MAX, on the other hand, presupposes that no true answer is strictly stronger than Alice laughed, thus excluding the alternatives *laughed(a⊕b)*, *laughed(a⊕c)*, and *laughed(a⊕b⊕c)*. This is illustrated in (11) by strikethrough text. Therefore, the inquiry-terminating approach correctly predicts that the sentence '*It was Alice who laughed*.' obtains the interpretation that Alice laughed and only Alice laughed.

Now consider the negated sentence in (12). In this case, MIN asserts that Alice did not laugh, nor did any of the pluralities containing Alice laugh. As Velleman et al. (2012: p. 455) write: "[S]ince laughed is a distributive predicate, this is just the same as saying Alice didn't laugh", shown in (13).

(12) It wasn't Alice who laughed.

$$\neg\text{MIN}(laughed(a)) \models \neg laughed(a) \wedge \neg laughed(a\oplus b) \wedge \neg laughed(a\oplus c) \wedge \neg laughed(a\oplus b\oplus c)$$

(13) $$\neg\text{MIN}(laughed(a)) = \neg laughed(a)$$

As for the MAX meaning component, illustrated in (14), it again presupposes that "no larger group including Alice laughed, though Alice alone might have" (Velleman et al. 2012: p. 455). Thus, the negated sentence in (12) asserts that Alice didn't laugh (MIN) and presupposes that no plurality containing Alice laughed (MAX).

(14) $$\text{MAX}(laughed(a)) \models \neg laughed(a\oplus b) \wedge \neg laughed(a\oplus c) \wedge \neg laughed(a\oplus b\oplus c)$$

Under the inquiry-terminating approach it is straightforward to explain why the series of sentences in (15) is predicted to be unacceptable. Namely, there is a contradiction given the distributive interpretation of the predicate *to laugh*. That is, the *it*-cleft in the first sentence asserts that Alice did not laugh, while the second sentence entails that Alice did laugh.

(15) #It wasn't Alice who laughed. Alice and Bob laughed.

a. *It wasn't Alice who laughed.*
Asserted (MIN): Alice did not laugh.
Presupposed (MAX): No plurality containing Alice laughed.
b. *Alice and Bob laughed.*
Asserted: Alice laughed.

Crucially, the situation is different when the predicate is interpreted non-distributively. Consider example (16) in a context such as (5b), in which there is a collective dish-washing event. Since under a non-distributive interpretation of the mixed predicate, the predicate predicates of the plurality but not of the atomic elements forming the plurality, it can be so that the predicate is false of the atomic individual *Kimberly* but true of the plurality *Kimberly and Helen*.

(16) [under the non-distributive interpretation in (5b):] Helen and Kimberly live together. Yesterday Helen cooked, but they did the dishes together.
It wasn't Kimberly who did the dishes. Kimberly and Helen (together) did the dishes.

a. *It wasn't Kimberly who did the dishes.*
Asserted (MIN): Kimberly did not do the dishes alone, although a plurality containing Kimberly may have done the dishes.
Presupposed (MAX): no entailment relationship between the alternatives; thus, no presupposition failure.
b. *Kimberly and Helen (together) did the dishes.*
Asserted: The plurality containing Kimberly and Helen did the dishes.

As a result, the sentences in (16) with a non-distributive interpretation result in a non-contradictory series of sentences, illustrated in (16a)–(16b). As for the presupposition, since there is no entailment between alternatives and thus no ordering with respect to one another, there is no presupposition failure and thus the series of sentences in (16) is predicted to be acceptable.[6]

In sum, under the alternative-based approach the acceptability of '*It wasn't* α *that did P.* α *and* β *did P.*' will depend on the interpretation of the predicate: while with distributively interpreted predicates the series of sentences is unacceptable, with non-distributively interpreted predicates it is acceptable.

6 That is, under a non-distributive interpretation of the predicate the alternative answers to the QUD are not ordered via entailment relations, and thus no alternative answer will ever be excluded by MAX.

3.2.4 Predictions

In this section, we spell out the different predictions for non-cleft SVO sentences and *it*-clefts depending on the distributive vs. non-distributive interpretation of the predicate.[7] We start with SVO sentences, an example of which is provided in (17).

(17) Kimberly didn't do the dishes. Kimberly and Helen did the dishes.

For the series of sentences in (17) with a distributive interpretation of the predicate, $x \notin [\![P]\!]$ followed by $(x \oplus y) \in [\![P]\!]$ will result in a contradiction. The contradiction arises because distributive predicates have atomic entities in their denotation, and therefore it follows from $(x \oplus y) \in [\![P]\!]$ that $x \in [\![P]\!]$, which contradicts the first sentence stating that $x \notin [\![P]\!]$. By contrast, for a non-distributive interpretation of the predicate, asserting $x \notin [\![P]\!]$ followed by $(x \oplus y) \in [\![P]\!]$ will not result in a contradiction, since collective predicates do not have atomic entities in their denotation, and thus it does not follow from $(x \oplus y) \in [\![P]\!]$ that $x \in [\![P]\!]$. Therefore, it is possible that $x \notin [\![P]\!]$ but $(x \oplus y) \in [\![P]\!]$.

Given this, non-cleft SVO sentences serve as a baseline measure for contradictions in distributive contexts and the lack thereof in non-distributive contexts. A summary of predictions is provided in (18).

(18) Kimberly didn't do the dishes. Kimberly and Helen did the dishes.

under the distributive interpretation in (5a):

a. 1st sentence: Kimberly did not do the dishes.

b. 2nd sentence: Kimberly did the dishes.
   $\rightsquigarrow$ contradiction $\Rightarrow$ **unacceptability** (indicated by × in Table 3.1)

under the non-distributive interpretation in (5b):

a. 1st sentence: Kimberly did not do the dishes alone, although a plurality containing Kimberly may have.

b. 2nd sentence: Kimberly and Helen did the dishes (but Kimberly alone did not do them).
   $\rightsquigarrow$ no contradiction $\Rightarrow$ **acceptability** (indicated by ✓ in Table 3.1)

Now consider an *it*-cleft sentence such as (19) and the predictions with distributively and non-distributively interpreted predicates.

(19) It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.

7 We are discussing here the inquiry-terminating approach for the sake of concreteness; however, the predictions extend to other analyses within the alternative-based approach to the semantics of clefts.

We will start with the homogeneity approach. Büring & Križ (2013: pp. 10–11) write:

> "In sum, **we believe that distributive and collective predicates behave identically in clefts**: If *a* is a proper part of those who *Q*, *it was* a *that Qed* is undefined, rather than false." [emphasis added]

They provide the following empirical generalization.

(20) #It wasn't Fred she invited. She invited Fred and Gord.
[ex. (3b) in Büring & Križ 2013]

Under the homogeneity approach in Križ 2017, the negated cleft sentence in (21) gives rise to the following truth-conditions (repeated from example (7) for the reader).

(21) It wasn't Kimberly who did the dishes.

true iff Kimberly does not overlap with the mereological sum of people who did the dishes
≈ *iff Kimberly did not participate in doing the dishes*

false iff Kimberly is identical to the mereological sum of people who did the dishes
≈ *iff Kimberly and only Kimberly did the dishes*

undef otherwise
≈ *Kimberly and Helen did the dishes*

Crucially, the truth-conditions of (21) do not change depending on the interpretation of the predicate. Therefore, in the context in which Kimberly and Helen did the dishes—whether they did it distributively as in (5a) or non-distributively as in (5b)—the homogeneity approach predicts the first sentence of the sequence in (22) to be undefined.[8]

(22) #It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.
under a distributive (5a) and non-distributive (5b) interpretation:
a. 1st sentence: undefined
⇝ undefinedness ⇒ **unacceptability** (indicated by × in Table 3.1)

In short, the homogeneity approach does not predict that the acceptability of '*It wasn't α that did P. α and β did P.*' will depend on the interpretation of the

8 Here we assume, following Križ 2017, that the negation in (22) is a truth-conditional negation. In Section 3.4 we will discuss the possibility of analyzing it as a metalinguistic negation, an option also mentioned in Büring & Križ 2013 and Križ 2017.

predicate: for both distributively and non-distributively interpreted predicates the series of sentences is expected to be unacceptable. Therefore, while the homogeneity approach predicts *it*-clefts to elicit equally unacceptable judgments under both predicate interpretations, non-cleft SVO sentences are predicted to be unacceptable with distributively interpreted predicates, on the one hand, yet acceptable with non-distributively interpreted predicates, on the other—thus, resulting in an interaction of predicate interpretation (distributive vs. non-distributive) and sentence type (*it*-cleft vs. SVO).[9]

By contrast, according to the inquiry-terminating approach in Velleman et al. 2012, *it*-clefts under distributively vs. non-distributively interpreted predicates will obtain divergent judgments. As Velleman et al. (2012: p. 455) write:

> We do [...] predict that [example (23)] should be infelicitous.
>
> (23) ??It wasn't Alice who laughed, it was Alice and Bob.
> [example (31) in Velleman et al. 2012]
>
> Note, though, that this prediction depends crucially on the fact that *laughed* is a distributive predicate. If we replace it with a non-distributive predicate, as in [example (24)], we predict felicity. [...] This matches our intuitions—though we must admit that these are subtle intuitions, and further empirical evidence here would be helpful.
>
> (24) It wasn't ALICE who moved the sofa, it was Alice AND the OTHER movers.
> [example (32) in Velleman et al. 2012]

Thus, the inquiry terminating approach makes the following predictions regarding the series of sentences in (25), which is parallel to non-cleft SVO sentences; cf. (18).

9 In previous literature it has been proposed that conjunction also gives rise to homogeneity (see Szabolcsi & Haddican 2004; but note that homogeneity is defined there differently from that in Križ 2017 and, moreover, it is defined only for distributive predicates). Nevertheless, Križ (2017: p. 23) writes: "[...] we wish to remain agnostic about how precisely conjunction and homogeneity interact. Our approach merely makes the general predication that the behaviour of conjunctions in clefts should align with how they interact with homogeneous predication in general." If we assume that conjunction does in fact give rise to homogeneity as defined in Križ 2017, however, then the SVO sentences '*Kimberly didn't do the dishes. Kimberly and Helen did the dishes.*' are predicted to give rise to a uniform pattern of responses in both distributive and non-distributive contexts. Yet, this is not what we found in our experimental data. We thank the anonymous reviewer who pushed us to clarify this issue.

(25) It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.

under the distributive interpretation in (5a):

a. 1st sentence
Asserted (MIN): Kimberly did not do the dishes.
Presupposed (MAX): No plurality containing Kimberly did the dishes.

b. 2nd sentence
Asserted: Kimberly did the dishes.
$\rightsquigarrow$ contradiction $\Rightarrow$ **unacceptability** (indicated by $\times$ in Table 3.1)

under the non-distributive interpretation in (5b):

a. 1st sentence
Asserted (MIN): Kimberly did not do the dishes alone, although a plurality containing Kimberly may have.
Presupposed (MAX): no entailment relationship between the alternatives; thus, no presupposition failure.

b. 2nd sentence
Asserted: The plurality containing Kimberly and Helen did the dishes (but Kimberly alone did not do them).
$\rightsquigarrow$ no contradiction $\Rightarrow$ **acceptability** (indicated by $\checkmark$ in Table 3.1)

In sum, the inquiry-terminating account in Velleman et al. 2012 predicts that the acceptability of '*It wasn't* α *that did P.* α *and* β *did P.*' will depend on the interpretation of the predicate—distributively interpreted predicates will be unacceptable, while non-distributively interpreted predicates will be acceptable—and the same predictions hold for the various other theories in the alternative-based approach. Moreover, *it*-clefts are predicted to show parallel response patterns to non-cleft SVO sentences, and thus no interaction of predicate interpretation and sentence type is expected.

A summary of predictions for the homogeneity approach and the alternative-based approach is presented in Table 3.1. In short, under the homogeneity approach acceptability for *it*-clefts and their non-cleft SVO counterparts is expected to show divergent patterns depending on the interpretation of the predicate: critically, *it*-clefts are predicted to be unacceptable across distributive vs. non-distributive interpretations, resulting in an interaction of predicate interpretation and sentence type. By contrast, under the alternative-based approach, *it*-clefts and their non-cleft SVO counterparts are predicted to show parallel response patterns, and thus no interaction. We tested these predictions in three experiments, which we turn to next.

	SVO	*it*-CLEFT	
		homogeneity	*alternative-based*
DISTRIBUTIVE	×	×	×
NON-DISTRIBUTIVE	✓	×	✓

Table 3.1 *Summary of predictions for the acceptability of non-cleft SVO and* it*-cleft sentences for the series of sentences such as* 'It's not α that did *P*. α and β did *P*.' *with distributively and non-distributively interpreted predicates.*

3.3 Experiments

In order to check the predictions in Table 3.1, we conducted three experimental studies using an acceptability judgment task with American English native speakers as participants. In Section 3.3.1 we present the methods and design for Experiment 1, in which written stimuli were used to test the predictions of the homogeneity and alternative-based approaches. Although we found a reliable main effect of distributive vs. non-distributive predicate interpretation on acceptability judgments, there remained a lot of variability in the response patterns for non-distributive interpretations of the predicates, which we postulate was mostly due to the use of written stimuli. Therefore, we conducted two follow-up studies, which we present in Section 3.3.2 and Section 3.3.3. The first follow-up study, Experiment 2, was based on Experiment 1, except auditory stimuli were used in order to control for the effects of prosody on the acceptability of the target sentences in context. The second follow-up study, Experiment 3, similarly employed auditory stimuli, but in this experiment we used solely mixed-predicates, unlike in the previous two experiments, in which we used both mixed and distributive predicates. Anticipating the results, in all three studies we found a robust effect of predicate interpretation on the acceptability of the sentences in context, for both SVO sentences and *it*-clefts, contrary to the predictions of the homogeneity approach but consistent with the alternative-based approach.

3.3.1 Experiment 1

3.3.1.1 Methods

Participants We tested 24 monolingual American English speakers (18 female, 6 male; mean age: 38, age range: 23–64).[10] All participants self-reported as having

10 For transparency, 26 participants in total completed the questionnaire but two participants were removed for not being monolingual speakers. Furthermore, we note that six of the participants saw a

grown up in the continental U.S.A. with English as their native language. All but two of the 24 participants reported speaking at least one foreign language, albeit with varying degrees of proficiency, and all participants had a Bachelor's degree or higher.

Materials Experiment 1 employed a fully-crossed 2x2 factorial design with the factors *Context* (2 levels: DISTRIBUTIVE, NON-DISTRIBUTIVE) and *Sentence* (2 levels: *it*-CLEFT, SVO). In total 40 target items and 80 filler items were tested, for a 1:2 target-to-filler ratio. All items had unique lexicalizations distributed in a Latin square design across four lists. For a full list of items, see Appendix C.2.

Crucially, we used predicates whose distributive vs. non-distributive interpretation was triggered by contextual manipulations rather than the lexical meaning of the predicate alone. For instance, the context in (26a) makes it clear that Carlos and Andrea biked individually, whereas (26b) establishes that the biking event applies to a collective plural entity. Contexts as in (26) were followed by a target sentence for participants to evaluate, which was either an *it*-cleft as in (27a) or a non-cleft SVO sentence as in (27b).

(26) *Context* [Exp. 1, Episodic, Item 10]

a. Carlos and Andrea like biking in a nearby forest. However, they have never seen each other there! Last week, Carlos biked on Monday and Andrea biked on Wednesday. [DISTRIBUTIVE]

b. Carlos and Andrea like biking. They own a tandem bike and they use it all the time! Last week, they biked in a nearby forest together. [NON-DISTRIBUTIVE]

(27) *Sentence*

a. It wasn't Carlos who biked. Carlos and Andrea biked. [*it*-CLEFT]

b. Carlos didn't bike. Carlos and Andrea biked. [SVO]

In this experiment both distributive and mixed predicates were used, the interpretations of which were manipulated contextually. Using distributive predicates was based on the observation in Renans 2016a,b that distributive predicates in clefts can be reinterpreted non-distributively as a rescue strategy. Thus, we assume that in the

version of the experiment which had a typo in two of the target items and a minor typo in the filler, which were corrected for the remaining participants. The typos and corrections were as follows: for target item 15 in habitual contexts, "Bill Lawrence" was corrected to "Bill and Lawrence"; for target item 06 in episodic contexts, "didn't won" was corrected to "didn't win"; and for filler item 07 "MoMa" was corrected to "MoMA". We found no difference in judgments despite the typos and have left the six participants in our data set.

context of (28b)—though not in (28a)—the predicate *to give birth* in (29) can be reinterpreted in a non-distributive manner.

(28) *Context* [Exp. 1, Episodic, Item 01]

a. Jacob and Ryan are Maria's children. Jacob is fifteen and Ryan is two. [DISTRIBUTIVE]

b. Jacob and Ryan are twins, sons of Maria. [NON-DISTRIBUTIVE]

(29) *Sentence*

a. It's not Ryan Maria gave birth to. She gave birth to Jacob and Ryan. [*it*-CLEFT]

b. Maria didn't give birth to Ryan. She gave birth to Jacob and Ryan. [SVO]

We thank the two anonymous reviewers who pointed out that the coerced interpretation of lexically distributive predicates in non-distributive contexts obtains the interpretation that the two individuals are participants in the same event at the same time or place, and neither was a participant in the event alone, giving rise to a spatio-temporally contiguous, communal, or coordinated interpretation rather than a 'truly' collective one.[11] For Experiment 1 and Experiment 2, in which we used both distributive as well as mixed predicates, we assume that after reinterpretation of the distributive predicate in a non-distributive manner there will no longer be an entailment relation between, e.g., '*Maria gave birth to Jacob and Ryan*', on the one hand, and '*Maria gave birth to Ryan*', on the other—for example, by coercing an

11 A similar observation was made by Onea (2007) regarding Hungarian pre-verbal focus sentences comparable to the stimuli here, such as in (ia)–(ib) (judgments from the original).

(i) a. ??Nem PÉTER kapott tízest, hanem Péter és PÀL (kapott tizest).
not Peter got ten-ACC but Peter and Paul got ten-ACC
'*It isn't Peter who got a ten (grade), it's Peter and Paul who got a ten (grade)*'

b. Nem Péter aludt a padlón, hanem Péter és Pál (aludt a padlón).
not Peter slept the floor.on but Peter and Paul slept the floor.on
'*It isn't Peter who slept on the floor; it's Peter and Paul who slept on the floor.*'
[ex. (12) and ex. (3) in Onea 2007, respectively, with minor modifications]

Onea (2007: 173) writes with regards to (ia): "[the sentence in (ia)] is weird for most speakers, except for some reading in which Peter and Paul got a grade for a joint work". Onea continues: "This shows that this kind of negation will only work in cases in which the conjunction delivered in the second clause can be conceived as referring to participants of the same event. Hence (ib) can only have the reading according to which Peter and Paul slept both on the floor at the same time. [. . .] But if Peter and Paul are the participants of a particular event, *the statement that Peter is the participant of the event is false*" (emphasis added). If this line of thinking is correct, then the results for Experiment 1 and Experiment 2 with lexically distributive predicates are unsurprising: it is clear from the context that neither individual was the sole participant in the relevant event.

interpretation of the predicate *to give birth* in the context of (28b) such that it only holds of twins. Crucially, in order to make sure that in non-distributive contexts a 'truly' collective non-distributive interpretation was obtained, for Experiment 3 we reran the experiment with mixed predicates only; see Section 3.3.3 for details.[12]

The decision to have contextual manipulations for predicate interpretation rather than using lexically distributive and collective predicates was threefold. First, as mentioned above, it has been observed that distributive predicates can be reinterpreted in a non-distributive manner as a rescue strategy (Renans 2016a,b). Thus, contextual manipulations were necessary to ensure that participants were presented with the interpretation we intended. Second, we aimed at having singular entities in the cleft pivot, both to keep the cognitive demands of the task to a minimum in terms of the number of referents introduced per trial as well as to follow the example sentences discussed in the cleft literature. Yet, singular entities are impossible with lexically collective predicates, such as *to disperse* or *to gather*. Third, contextual manipulations allowed us to present participants with the intended interpretation of the predicate without changing the predicates themselves, thus making a fully-crossed 2x2 design for the factors *Context* x *Sentence* possible.

There are two additional points to make regarding the construction of the target items in Experiment 1: First, we also systematically manipulated aspectual reference to ensure that the decisive factor for the evaluation of the sentence is the distributive vs. non-distributive interpretation of the predicate, and not aspectual interpretation. Thus, there was an additional between-item factor *Aspect* (2 levels: EPISODIC, HABITUAL), such that all conditions were tested with both episodic (20 items) and habitual (20 items) readings; see Appendix C.2 for details.

Second, in order to make sure that there was no influence of the order of conjuncts on acceptability judgments, for the factor *Sentence* we additionally counterbalanced the ordering of the conjuncts in the second of the two sentences. The even-numbered items (2, 4, 6, and so on) for the episodic readings and the odd-numbered items (1, 3, 5, and so on) for the habitual readings had the word order '… α *and* β *did P*' in the second sentence, as in the examples (27b)–(27a) above (e.g., '*It wasn't Carlos who biked. Carlos and Andrea biked.*'). By contrast, the odd-numbered items (1, 3,

12 As one reviewer pointed out, the predictions for Velleman et al. 2012 differ when assuming that an entailment relation holds despite a spatio-temporally contiguous, communal, or coordinated interpretation: acceptability judgments for 'truly' lexical predicates, with reference to entailment, should be the same across contextual manipulations, contrary to fact. In this case, the results for Experiment 1 and Experiment 2 are potentially unexpected for all theories discussed here ("potentially" since mixed predicates were also used). We thank the reviewer for pointing this out. Given the empirical picture presented here—in particular the results of the follow-up Experiment 3, in which only mixed predicates were used—we will leave the issue of entailment relations for 'truly' collective vs. spatio-temporally contiguous, communal, or coordinated interpretations as a compelling puzzle for future research.

5, and so on) for the episodic readings and the even-numbered items (2, 4, 6, and so on) for the habitual readings had the word order '… β *and* α *did P*' in the second sentence (e.g., '*It's not Patricia who plays computer games. Martha and Patricia play computer games.*'). There is no evidence that the order of conjuncts influenced acceptability judgments.

Regarding the filler items, we included multiple cleft and cleft-like structures, both with and without negation (see Appendix C.3 for sample filler items). Specifically, we included 20 *wh*-clefts, 20 expletive sentences, 20 *it*-clefts, and 20 definite pseudoclefts. An example trial for the *wh*-cleft is as follows (example filler item F01).

(30) *Context*: Diana is spending her holidays in California and Gary is spending his holidays in Texas.
Filler: Where Diana and Gary are spending their holidays is not Canada. They're spending their holidays in the USA. [F01]

Of the 20 *wh*-clefts, 10 included negation (as in the above), whereas 10 did not (for example items, see F01–F04 in the appendix). Moreover, 5 of the sentences with negation were intended to be semantically coherent and acceptable (as in the above) and 5 semantically incoherent and unacceptable (see, e.g., F02); the same held for the sentences without negation (see, e.g., F03–F04).

With the term expletive sentences we intend sentences which at the onset are the same as *it*-clefts, such as the following.

(31) *Context*: George gave a radio interview in which he recommended two museums to visit: MoMA in New York and The Louvre in Paris.
Filler: It's obvious that George recommended MoMA and The Louvre. [F07]

Again, of the 20 expletive sentences, 10 included negation, whereas 10 did not (as in the above; for further example items, see F05–F08 in the appendix). Moreover, 5 of the sentences with negation were intended to be semantically coherent and acceptable (as in the above) and 5 semantically incoherent and unacceptable (see, e.g., F06); again, the same held for the sentences without negation (see, e.g., F07–F08).

Finally, the *it*-cleft and definite pseudocleft manipulations were such that the element which appeared in the cleft pivot or after the copular verb included expected or unsurprising entities in half the trials (e.g., *photo* in the context of social media) and unexpected or surprising entities in half the trials (e.g., *ransom note* in the same context). In all trials, a violation of exhaustivity occurred in the second conjunct. These items were distributed in a Latin square design across the four lists. Examples for the *it*-cleft (F09–F10) and definite pseudocleft (F11–F12) trials are as follows.

(32) *Context*: Michael is on his favorite social network each and every day.
Filler: It's a {photo/ransom note} that Michael posted and he posted a video. [F09–F10]
Filler: The thing that Michael posted is a {photo/ransom note} and he posted a video. [F11–F12]

Procedure In each trial, participants were presented with a short description of the context and a target sentence, both of which were in written form. They were instructed that their task was to provide judgments on a scale from 1 ('unacceptable') to 7 ('acceptable') of the sentences in context. In order to become familiar with the task, participants were given three practice trials before the experiment began. The experiment was conducted online using the free software platform OnExp (GNU General Public License) hosted at the Universität Göttingen (https://onexp.textstrukturen.uni-goettingen.de/). All items were randomized during presentation. The task took about 35–45 minutes to complete, and participants were compensated $7.00 for their participation.

3.3.1.2 Results

We conducted a Bayesian ordinal mixed logistic regression analysis; see Bürkner & Vuorre 2018 for a tutorial on ordinal regression using *brms*, as well as Liddell & Kruschke 2018 for a discussion of ordered-probit models in a Bayesian framework and Kruschke & Liddell 2018 for general introductions to Bayesian modelling. We used the statistics software *R* (v. 3.5.2, GPL-2 | GPL-3; R Core Team 2019) with the *brms* package (v. 2.7.0, GPL >= 3; Bürkner 2017, 2018), which provides an interface to fit Bayesian models using Stan (New BSD License; Stan Development Team 2018).

We used sum contrasts for the factors *Sentence* (CLEFT –1, SVO 1) and *Context* (DISTRIBUTIVE –1, NON-DISTRIBUTIVE 1), and we included maximal random-effects structures in our statistical models, with varying intercepts and slopes for both participants and items. Moreover, we used regularizing, weakly-informative priors in order to downweight extreme values and obtain stable inferences (Vasishth et al. 2018).[13] We report point estimates of the parameters from the posterior distribution

13 The model for Experiment 1, which included the factor *Aspect* (EPISODIC –1, HABITUAL 1), is as follows; note that the models for Experiment 2 and Experiment 3 were identical minus *Aspect*.

```
brm(formula = Acceptability ~ 1 + Context * Sentence * Aspect +
              (1 + Context * Sentence * Aspect | Participant) +
              (1 + Context * Sentence | Item),
    data = exp1, family = cumulative('probit'),
    prior = c(set_prior('normal(0, 3)', class = 'Intercept'),
```

along with the 95% credible intervals (given these data and the model, the interval containing the 95% most credible values of the parameter, abbreviated as 95% CrI). If the credible interval overlaps with zero we interpret that as lack of evidence of an effect: we remain uncertain whether the parameter is zero (no effect) or has the wrong sign. In cases when the overlap is very minimal, however, with almost all of the probability density on one side of zero, we will report this as weak evidence that an effect is present. By contrast, should the credible interval have no overlap with zero we will interpret that as reliable evidence of a robust effect.

To get a sense of the shape of the data for Experiment 1, we start with visual inspection of the histogram in Figure 3.1 showing the frequency of the discrete acceptability ratings from '1'–'7'. As can be seen, sentences in contexts triggering a distributive reading were largely judged as unacceptable, with ca. 92% (each: 221/240) of acceptability judgments for both clefts and non-cleft SVO sentences falling below the middle '4' value, with a clear majority of judgments at '1', the lowest rating on the scale. In comparison, contexts triggering a non-distributive reading—though having a large number of negative responses—nonetheless display a wide spread across the scale: 58% (140/240) of the acceptability judgments for clefts and 64% (153/240) for SVO sentences fell below the middle '4' value, whereas 28% (67/240) of the judgments for clefts and 27% (61/240) for SVO sentences fell above the middle '4' value.

Nevertheless, statistically—given the highly positive coefficient of *Context* from the ordered-probit model ($\hat{\beta}$ = 0.77; 95% CrI: 0.53, 1.01)—results indicate that there was a main effect of context: sentences in distributive contexts are reliably judged as less acceptable than sentences in non-distributive contexts. Moreover, there is no statistical evidence of an effect for either *Sentence* ($\hat{\beta}$ = –0.05; 95% CrI: –0.19, 0.07) nor evidence of an interaction of *Context* x *Sentence* ($\hat{\beta}$ = 0.01; 95% CrI: –0.10, 0.12), as the credible intervals for both estimates have a high degree of overlap with zero.

```
              set_prior('normal(0, 3)', class = 'b'),
              set_prior('normal(0, 3)', class = 'sd'),
              set_prior('lkj(2)', class = 'cor')),
      inits = 0, iter = 4000, cores = 4, chains = 4,
      seed = 2701, control = list(adapt_delta = 0.99))
```

We note that there was some weak evidence of an effect of *Aspect* ($\hat{\beta}$ = –0.13; 95% CrI: –0.29, 0.02) and of an interaction of *Aspect* x *Context* ($\hat{\beta}$ = –0.11; 95% CrI: –0.24, 0.03); nonetheless, the credible intervals suggest that the estimates may in fact be zero or on the other side of zero. Assuming for now a true effect exists, we think it is plausible that—as a rescue strategy—it may be easier to coerce a non-distributive interpretation of the predicate in episodic contexts than in habitual ones, since the latter leave more space for additional interpretation. For instance, participants may ask themselves whether the agents always really did the relevant activity together. That said, statistically the effect of aspect reported here is weak at best and nevertheless unreliable, and we will not discuss this further.

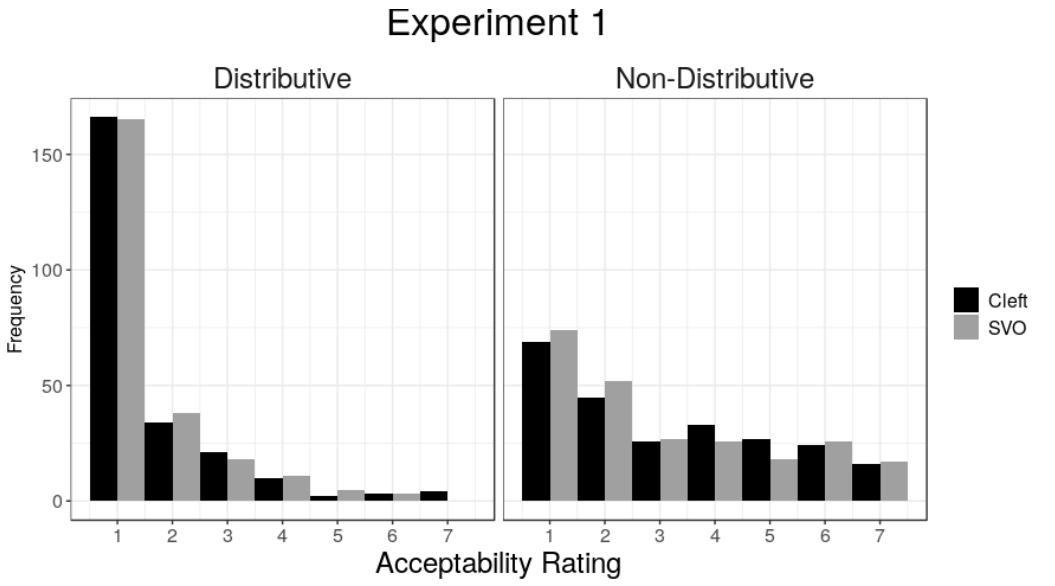


Figure 3.1 *Acceptability ratings (7-point scale: 1* 'unacceptable' *– 7* 'acceptable'*) shown as histograms for Experiment 1 (written).*

Interim discussion The main result of the first study is that the interpretation of the predicate was found to influence the acceptability of both *it*-clefts and SVO sentences: that is, sentences in contexts triggering a distributive interpretation were judged as overall less acceptable than sentences in contexts triggering a non-distributive interpretation, with no evidence of an interaction with sentence type. The parallel response patterns for SVO sentences and clefts is consistent with the alternative-based approach; furthermore, the results pose a direct challenge for the homogeneity approach, since *it*-clefts are predicted to be equally unacceptable under both distributive and non-distributive interpretations, contrary to what we found.

That said, although we found a difference between contexts as expected in the alternative-based approach, the ratings for both SVO and cleft sentences under a non-distributive interpretation were spread broadly across the scale—and included a high frequency of low judgments, which was not predicted. We can think of two reasons that might have caused the relatively low acceptability ratings in non-distributive contexts. First, the low judgments could be a local effect (cf., for instance, Hemforth 2018). Namely, there were twice as many filler items as target items in Experiment 1, many of which were intended to be perfectly coherent and acceptable; see the discussion of the filler items. By contrast, the judgments for clefts and SVO sentences are quite subtle. In fact, four of the conditions in the filler (namely, *wh*-clefts and

expletive sentences with and without negation) were judged as acceptable by participants with 75% to 96% of the judgments above the middle '4', of which a majority were '7', the highest rating on the scale. The almost ceiling-like responses for these conditions could in turn push the evaluation of the other less-robust conditions lower.

Second, and crucial for the two follow-up studies presented in the following sections, the overall low judgments could be caused by the varying implicit prosody participants assigned to the written stimuli (Fodor 2002a,b, Koizumi 2009). That is, contrastive focus—and not simple declarative prosody with H* pitch accent at the beginning of the sentence and a falling boundary tone—appears to be important to make the sentences in collective contexts acceptable.[14] Thus, in order to control for prosodic assignment in our stimuli we ran two follow-up studies using auditory stimuli.

3.3.2 Experiment 2

In this and the following section we discuss two follow-up experiments using auditory stimuli instead of written stimuli.

3.3.2.1 Methods

Participants For Experiment 2, we tested 32 monolingual American English native speakers (12 female, 20 male; mean age: 30, age range: 18–47). All participants were self-reported American English native speakers who grew up in the continental U.S.A. Of these participants, 27 reported knowing no foreign languages, while 5 reported speaking at least one foreign language; as for education, 21 participants had a Bachelor's degree or higher, 10 had a high school degree, and 1 had not completed high school.

Materials Only a subset of the stimuli from Experiment 1 was used for the follow-up Experiment 2. Doing so allowed us to reduce the length and complexity of the experiment given the more time-consuming and cognitively-demanding task of listening to and processing auditory stimuli. Thus, in contrast to the experiment with written stimuli, Experiment 2 tested a total of 20 target items (all with an episodic interpretation) plus 32 filler items, the latter being a subset of the *wh*-cleft and expletive sentence filler items from Experiment 1. All items were presented with a unique lexicalization in each condition, distributed in a Latin square design across four lists.

Although the filler and target stimuli in Experiment 2 were a subset of the stimuli in Experiment 1, note that for the factor *Sentence* we changed the order of conjuncts

14 We are thankful to the editor for bringing this to our attention.

in the second clause in the target items. That is, whereas in the written version of the experiment the odd- and even-numbered items differed (see Section 3.3.1 "Material" for details), in the auditory version of the experiment all items followed the same pattern: the argument which appeared in the cleft pivot was invariably first amongst the two conjuncts and the conjunction remained unstressed (i.e., '… α *and* β *did P.*') (cf. Križ 2017: p. 23 regarding stressed vs. unstressed conjunction). This was done in order to keep the prosodic contours consistent across the items; see Appendix C.2.1.

A male native English speaker from Canada in his mid-20s recorded all of the target and filler stimuli in a sound-proof acoustic lab with the audio-editing software Audacity (v. 2.2.2, GPL-2; Audacity Team 2018). To assist the speaker while reading aloud during the recording session, we provided a printout of the target sentences with pitch accents indicated by the use of all capitals on the stressed syllable, e.g., '*It wasn't CARlos who biked. Carlos and AnDREA biked*'. By contrast, for the filler sentences the instructions to the speaker were to read the sentences out loud in a way that felt natural. All sentences were recorded at least twice, from which we—with help from our student research assistant—selected the best recording for the experiment based on two criteria: (i) the pitch accent was placed at the intended location, and (ii) the recording sounded natural. Pitch accents in the target sentences were evaluated by the research assistant using the phonetics software Praat (v. 6.0.37, GPL-2; Boersma & Weenink 2018, Boersma 2001).[15]

Procedure The procedure for Experiment 2 was identical to Experiment 1, with one major difference: although the context was still presented in written form, the target stimuli were now presented in auditory form. The instructions were identical to the written version of the experiment with the exception that we also provided instructions regarding, e.g., adjusting the volume of the participants' headphones; additionally, we used auditory stimuli for the practice trials.

As before, the experiment was conducted online using OnExp. Participants were recruited and payed via Prolific (https://www.prolific.ac/). Again, all items were randomized during presentation. Given the reduced length of the experiment, participants were compensated ca. $5.50 for their participation in Experiment 2.

3.3.2.2 Results

Again, we start with visual inspection of Figure 3.2 showing the histogram of acceptability ratings in order to get a sense of the overall shape of the data. When controlling for prosody, one sees that sentences in contexts triggering a non-distributive reading

15 With permission, all recordings of the target stimuli are available at https://gitup.uni-potsdam.de/deveaugh/it-clefts-collective-distributive/tree/master/auditory-stimuli.

were judged as generally acceptable compared to sentences in contexts triggering a distributive reading, which were overall judged as unacceptable. Specifically, the percentage of '5'–'7' ratings in non-distributive contexts was ca. 72% for both clefts (115/160) and SVO sentences (116/160), with the majority of judgments at '6'–'7'; compare that to the percentage of '1'–'3' ratings at 19% (each: 31/160) for both sentence types, contrasting with the results of Experiment 1 with written stimuli. In distributive contexts, the opposite pattern was found: most of the judgments were at the low end of the scale, with the percentage of '1'–'3' ratings at 56% (89/160) for clefts and 62% (99/160) for SVO sentences; cf. '5'–'7' ratings at 28% (45/160) for clefts and at 24% (38/160) for SVO sentences.

Just as for Experiment 1, we conducted a Bayesian ordinal mixed logistic regression analysis using sum contrasts for *Sentence* (CLEFT –1, SVO 1) and *Context* (DISTRIBUTIVE –1, NON-DISTRIBUTIVE 1). The highly positive coefficient estimate for *Context* ($\hat{\beta}$ = 0.95; 95% CrI: 0.62, 1.29) indicates that participants judged sentences in contexts triggering a distributive reading as less acceptable than sentences in contexts triggering a non-distributive reading—a statistically robust effect replicating the results of Experiment 1. Unlike Experiment 1, however, there is also weak evidence of an effect of *Sentence* ($\hat{\beta}$ = –0.09; 95% CrI: –0.20, 0.02), with non-cleft SVO sentences being judged as less acceptable overall than their *it*-cleft counterparts. Finally, we found no evidence of an interaction of *Context* x *Sentence* ($\hat{\beta}$ = 0.02; 95% CrI: –0.10, 0.15).

Interim discussion As in Experiment 1, the results in Experiment 2 show that the distributive vs. non-distributive interpretation of the predicate had an influence on acceptability. As before, sentences in contexts triggering a distributive interpretation were judged as less acceptable than sentences in contexts triggering a non-distributive interpretation, and no interaction with sentence type was found. Importantly, once we controlled for the prosodic contour of the items in the auditory version of the experiment, the results became more clear in the directionality on the ordinal scale, with non-distributive contexts eliciting mostly acceptable judgments and distributive contexts eliciting mostly unacceptable judgments. In the discussion of Experiment 3, we will return to the result that non-cleft SVO sentences were judged slightly worse than *it*-clefts, since there we replicate the effect of sentence type found in Experiment 2.

One worry that the anonymous reviewers had was that in Experiment 1 and Experiment 2 both mixed and distributive predicates were used. Therefore, we ran a third follow-up experiment (Experiment 3) with mixed predicates only, which we turn to next.

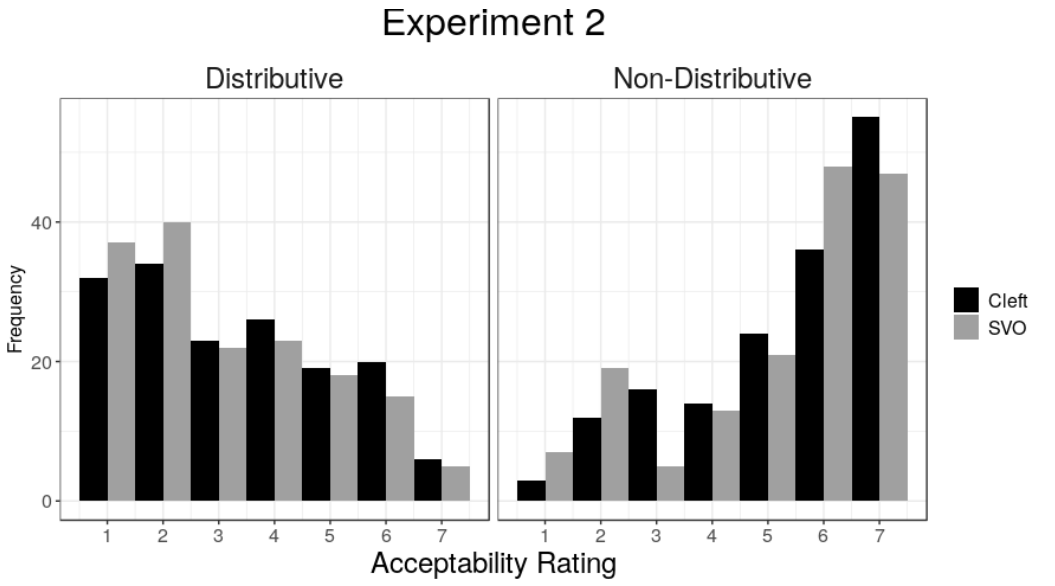


Figure 3.2 *Acceptability ratings (7-point scale: 1 'unacceptable' – 7 'acceptable') shown as histograms for Experiment 2 (auditory).*

3.3.3 Experiment 3

The methods and design in Experiment 3 were the same as in Experiment 2, and therefore we will focus here on the crucial differences between the experiments, in particular, the predicates used in the target stimuli.

3.3.3.1 Methods

Participants For Experiment 3, 40 monolingual American English native speakers (20 female, 16 male, 3 other, 1 no answer; mean age: 31, age range: 18–59) completed the task. All participants self-reported as having grown up in the continental U.S.A. Of the 40 participants, 30 reported knowing no foreign languages, while 10 spoke at least one foreign language. As for education, 25 participants had a Bachelor's degree or higher, 15 had a high school degree, and 1 participant did not respond to the question.

Materials The second auditory follow-up, Experiment 3, was the same as Experiment 2 with one crucial difference: we used mixed predicates only. That is, whereas in the first two experiments we included target items which are considered truly distributive—assuming they can be coerced into a non-distributive interpretation

as a rescue strategy; see Section 3.2 and Section 3.3.1.1 under "Materials"—for Experiment 3 we had solely mixed predicates for the target items. Furthermore, all target trials were controlled for to ensure that the interpretations of the predicates in non-distributive contexts were truly collective and not just spatio-temporally contiguous, communal, or coordinated (see, e.g., Lasersohn 1998, Syrett & Musolino 2013). The diagnostics for being considered truly collective are as follows.[16]

i. *Separate-conjuncts test*: if the sentence can be rephrased as two separate conjuncts, the predicate is not really collective. Example: *Madison won the marathon, and Abigail (also) won the marathon.*

ii. *Time-locked test (via "at the same time")*: if the event can be time-locked with a phrase such as "at the same time", the predicate is not really collective. Example: *Owen proposed to Alice and Linda at the same time.*

iii. *"Each" test*: if the sentence can be rephrased with *each*, the predicate is not really collective. Example: *Benjamin and Anne each gave a speech.*

iv. *Just-one-individual test*: if the sentence can be felicitously rephrased such that just one of the individuals has the property in question, the predicate is not really collective. Example: *Of a collective piano-lifting event where A and B work together to lift the piano and move it from one room to another room. Q: Did A (single-handedly) lift the piano? A: No!*

Target items which satisfied the above tests were reused from Experiment 2. All remaining target items were constructed *ex novo*, although we tried to use predicates which have been discussed in the previous experimental literature as being mixed; see Appendix C.4 for a full list of target stimuli from Experiment 3.

With the exception of the predicates used in the target stimuli, Experiment 2 and Experiment 3 were identical. Moreover, the same native speaker who was recorded for the auditory stimuli in Experiment 2 was again recorded for the target items in Experiment 3; see Section 3.3.2.1 under "Materials". We note that, whereas all target items were newly recorded for Experiment 3 (including target items which were the same as in early versions of the experiment), for the filler items we reused the recordings from Experiment 2. Thus, in order to correct minor differences in volume levels, the new recordings had to be adjusted slightly using the audio-editing software Audacity.

16 We thank the anonymous reviewers for raising this issue and the editor for providing us with the range of tests described here.

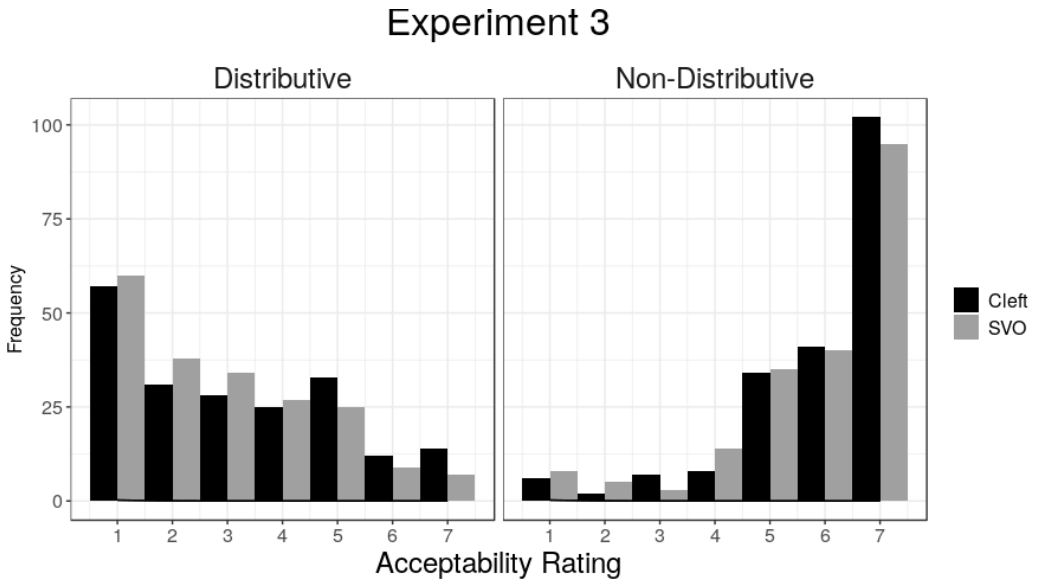


Figure 3.3 *Acceptability ratings (7-point scale: 1* 'unacceptable' *– 7* 'acceptable'*) shown as histograms for Experiment 3 (auditory, mixed predicates only).*

Procedure The procedure for Experiment 3 was identical to Experiment 2, although participants were now compensated ca. $6.50 in Experiment 3 for the 20–30-minute task.[17]

3.3.3.2 Results

As seen in the histogram in Figure 3.3 for Experiment 3, we find the same general division of acceptability ratings as in Experiment 2: in non-distributive contexts the percentage of acceptable '5'–'7' ratings is 89% (177/200) for clefts and 85% (170/200) for SVO sentences, with a majority of judgments at '6'–'7'; cf. about 8% '1'–'3' ratings for clefts (15/200) and SVO sentences (16/200). By comparison, the percentage of less acceptable '1'–'3' ratings in distributive contexts was 58% (116/200) for clefts and 66% (132/200) for SVO sentences; cf. '5'–'7' ratings at 30% (59/200) for clefts and 21% (41/200) for SVO sentences.

17 We decided to raise payment from £4.00 in Experiment 2 to £5.00 in Experiment 3 (the Prolific interface uses British pounds, which we converted into US dollars for presentational purposes above). We did so after becoming aware of issues related to unpaid work on crowdsourcing platforms given that participants must "log in to the site, answer a plethora of screening questions, locate a survey for which one is qualified" before they may begin the task, as described in the report on fair pay in digital labor platforms from the United Nation's *International Labour Organization* (Berg et al. 2018: p. 53).

As before, we fit a Bayesian ordinal mixed-effect logistic regression model with sum contrasts for *Sentence* (CLEFT –1, SVO 1) and *Context* (DISTRIBUTIVE –1, NON-DISTRIBUTIVE 1). Experiment 3 replicated the results of the previous two experiments. We found a highly positive coefficient estimate for *Context* ($\hat{\beta}$ = 1.54; 95% CrI: 1.23, 1.87), with participants judging sentences in distributive context as less acceptable than sentences in non-distributive contexts, and this effect was robust. Moreover, there was a reliable negative effect of *Sentence* ($\hat{\beta}$ = –0.12; 95% CrI: –0.24, –0.01), with SVO sentences judged as slightly less acceptable than *it*-clefts. However, we again failed to find any evidence of an interaction of *Context* x *Sentence* ($\hat{\beta}$ = 0.04; 95% CrI: –0.07, 0.15).

3.3.4 Interim discussion

To sum up all three studies: The results of Experiment 2 and Experiment 3 replicate statistically the results of Experiment 1. Moreover, by using auditory stimuli (Experiments 2 and 3) as well as having mixed-predicates only (Experiment 3) the results became ever more clear: there was a reliable and robust effect of predicate interpretation, in that contexts triggering a distributive interpretation had lower rates of acceptability than contexts triggering a non-distributive interpretation. Finally, no indication of an interaction with sentence type was found.

As an anonymous reviewer pointed out, while the results of Experiment 3 indeed constitute good evidence against the homogeneity approach, the results of Experiment 1 and Experiment 2 potentially remain puzzling due to the use of lexically distributive predicts ("potentially" since mixed predicates were also used). The puzzle is due to the possible entailment relation with truly distributive predicates even with a spatio-temporally contiguous, communal, or coordinated interpretation, as discussed in Section 3.3.1 under "Materials" (see, in particular, Footnote 12). For a series of sentences similar to the one used here, Renans (2016a,b) showed that for *nì*-clefts in Ga (Kwa), an under-researched language spoken in Ghana, distributive predicates, such as *to give birth*, can be reinterpreted in a non-distributive manner as a rescue strategy. For instance, example (33) is claimed to be unacceptable—unless Kofi and Emmanuel are twins (Renans 2016a: Fn. 43).

(33) #Jèèè Kòfí nì Màríà fɔ́. È-fɔ́ Kòfí kɛ̀ Emmanuel.
NEG Kofi PRT Maria give.birth 3SG-give.birth Kofi and Emmanuel
'*It's not Kofi who Maria gave birth to. She gave birth to Kofi and Emmanuel.*' [ex. (114) in Renans 2016a]

Along these lines, we had assumed that coercing a non-distributive interpretation of an otherwise lexically distributive predicate in the first two experiments would result in no entailment, for instance, between a sentence such as '*Maria have birth to Jacob*

and Ryan' and the sentence '*Maria gave birth to Ryan*', in cases when the predicate *to give birth* is understood as applying to twins. Saying that, the more robust results in Experiment 3 could be due to the fact that only mixed predicates were used, unlike in Experiments 1 and 2, in which both mixed and lexically distributive predicates were used. On the other hand, if for truly distributive predicates an entailment relation remains despite obtaining a non-distributive interpretation (spatio-temporally contiguous, etc.), then the results of Experiment 1 and Experiment 2 reported here are potentially unexpected for both homogeneity and the alternative-based account, an interesting issue we leave for future research.

As in Experiment 2, we found an effect of sentence type in Experiment 3: non-cleft SVO sentences were judged as slightly worse than *it*-clefts overall. Although a discussion on the discourse conditions of clefts is beyond the scope of this paper, this finding appears to be in line with several claims found in the literature: it has been argued that clefts encode a stronger degree of contrastiveness or contrariness than their canonical non-cleft counterparts (Byram-Washburn, Kaiser & Zubizarreta 2013, Destruel & Velleman 2014, Destruel et al. 2015, Destruel, Beaver & Coppock 2017). In this light, since the series of sentences tested here had a strong contrastive flavor, it is unsurprising that participants found clefts to be (at least slightly) better than SVO sentences.

In sum, the results of the three experiments reported here pose a direct challenge for the homogeneity accounts of cleft exhaustivity. For the series of sentences '*It's not* α *that did P.* α *and* β *did P.*' the homogeneity approach predicts *it*-clefts to be equally unacceptable under both a distributive and a non-distributive interpretation, contrary to what we found.[18] Furthermore, we found no evidence of a difference between *it*-clefts and SVO sentences in distributive and non-distributive contexts, a finding which is consistent with the alternative-based approach.

18 It was suggested by an anonymous reviewer that the experimental data from Experiments 1 through 3 can be accounted for under Büring & Križ's (2013) account of homogeneity if we assume that in non-distributive contexts—but not in distributive ones—participants interpreted the first sentence distributively and the second one non-distributively. The reviewer suggested that, in that case, a series of sentences such as '*It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.*' would turn out to be true and thus acceptable. This, in turn, could explain why sentences in non-distributive contexts were rated better than the same sentences in distributive contexts. We can see, however, a couple of problems with this approach. First, it is unclear why participants would interpret one and the same predicate once distributively and once non-distributively in a context that specifies for a non-distributive interpretation of the predicate. Second, even if the first sentence is interpreted distributively, and hence the assertion is true (i.e., Kimberly alone did not do the dishes), under Büring & Križ's (2013) account it still gives rise to the presupposition that Kimberly is not a proper mereological part of the sum of people who did the dishes, which clashes with the second sentence stating that Kimberly *is* a proper mereological part of the sum of people who did the dishes. Thus, this series of sentences is predicted to be unacceptable in the end. We thank the reviewer for asking us for clarifications on this issue.

3.4 Discussion

The results reported here suggest that the series of sentences '*It's not α that did P. α and β did P.*' is more acceptable in contexts triggering non-distributive interpretations than in contexts triggering distributive interpretations. Although the literature has noted that intuitions for such sentences are quite subtle—to the extent that Velleman et al. (2012: p. 455) explicitly state "further empirical evidence [...] would be helpful"—our findings suggest that the differences in interpretation are robust: the effect was replicated across three studies, using various modalities (written vs. auditory) and predicates (Experiment 1 & 2 vs. Experiment 3).

Although these results are a challenge for the homogeneity approach, there is one issue that remains to be addressed: the role of negation.[19] Throughout the discussion, we have assumed that the negation in the experimental items is truth-conditional negation; that is, that it targets the asserted meaning component. At the same time, however, Büring & Križ (2013: Fn. 1) and Križ (2017: Fn. 1)—following Horn (1981, 1985, 1989)—admit that the series of sentences '*It's not α that did P. α and β did P.*' can sometimes be accepted if the negation in the first sentence is not a truth-conditional negation but a metalinguistic negation. We would like to discuss this issue in detail here.

It has been observed in the literature that in some cases negation does not target the assertion but rather some other non-asserted meaning component of a sentence (see, for example, Karttunen & Peters 1979, Horn 1985, 1989, Guerts 1998). Consider, for instance, example (34).

(34) The king of France is not bald.
   a. PRESUPPOSITION: *there is a king of France*
   b. ASSERTION: *the king of France is not bald*

The sentence in (34) can obtain either an interpretation in which the presupposition projects and only the assertion is negated, as in (35a), or an interpretation in which both the presupposition and the assertion are negated, as in (35b). One way of formalizing this is by referring to a global vs. local accommodation of the presupposition; see, e.g., Beaver & Zeevat 2007, von Fintel 2008, Romoli & Sauerland 2017. Concretely, (35a) illustrates when the presupposition is accommodated globally, i.e., at the sentential level. On the other hand, (35b) illustrates when the presupposition is accommodated locally, under the scope of negation.

(35) a. There is a king of France and he is not bald. (*global acc.*)
   b. There is no king of France and he is not bald. (*local acc.*)

19 We thank the anonymous reviewer who pushed us to discuss this issue in detail.

Regarding clefts, Horn (1989) claims that examples such as (36a) and (36b) are acceptable with metalinguistic negation, which targets the non-asserted inferences that Mary kissed nobody other than John and ate nothing other than pizza, respectively.[20]

(36) a. It wasn't John that Mary kissed—it was John and Bill.
b. It wasn't a pizza that Mary ate—it was a pizza, a calzone, and a side of ziti. [exs. (46c)–(46d) in Horn 1989]

Now, the crucial point here is that this account can only work if clefts give rise to at least two meaning components—an asserted and non-asserted meaning—which the (metalinguistic) negation can target. As discussed in Sections 3.2.2 and 3.2.3, it is indeed the case that there are multiple meaning components in the alternative-based approach to cleft exhaustivity, but, crucially, not in the homogeneity approach. In Križ's (2017) formulation, homogeneity is neither a presupposition nor an implicature, but a way of capturing the trivalent logic characterizing the truth-conditions of sentences with homogeneous predicates. Since homogeneity is not modelled as another layer of meaning in clefts (e.g., it is neither a presupposition nor an implicature), the account of Horn (1989) is difficult to applicate here.

One way of doing this was suggested to us by an anonymous reviewer. The reviewer proposed that the metalinguistic negation in the cleft structure in (36a) may not be a negation of the cleft in (37), but its non-cleft SVO counterpart in (38) with focus on John—which gives rise to the implicature that Mary kissed John and nobody else, shown in (38a). The role of the metalinguistic negation in the cleft in that case would be to cancel this implicature, shown in (38b) (cf. Horn 1989: §6.6).

(37) It was John that Mary kissed.

(38) Mary kissed [John]$_F$.
a. (*implicature*) Mary kissed only John.
b. (*metalinguistic negation*) It is not the case that Mary kissed only John.

We share the reviewer's worry, though, that even if the metalinguistic negation is applicable to these cases, it still cannot account for the distinction between distributive vs. non-distributive contexts found in all three experiments.

By contrast, homogeneity is in fact modeled as another layer of meaning in Büring & Križ's (2013) formulation, but Križ (2017: p. 18) rejects this proposal, stating: "[Büring & Križ (2013)] say explicitly that they [model homogeneity as a presupposition] merely for lack of alternatives and do not want to strongly commit to a particular status of the neither-truth-nor-falsity that is observed with definite

20 Note, however, that in light of the results of our three experiments, the empirical generalizations provided in Horn (1989) are not quite sufficient. Yes, (36a) and (36b) are acceptable, but only if the distributive predicates *to kiss* and *to eat* are interpreted in a non-distributive manner.

descriptions and cleft sentences [...]". Nevertheless, keeping the formulation as in Büring & Križ 2013 instead and treating homogeneity as a presupposition would allow one to adopt Horn's (1989) analysis of metalinguistic negation for clefts, potentially accounting for our results while maintaining (a version of) the homogeneity approach. As Križ (2017: §4.3) points out, however, this formulation is ill-defined for complex sentences such as clefts with definite descriptions in pivot position, for which reason Križ proposed an update of Büring & Križ's (2013) account.[21]

Another possibility is that metalinguistic negation signals instead that the wrong form was used while keeping the assertion intact. This would be similar to the correction of the pronunciation in (39), which Krifka (2008: p. 248) refers to as *expression focus*.

(39) John didn't come to BERlin. He came to BerLIN.

In the case of (39), the metalinguistic negation does not negate the assertion that John came to Berlin but the fact that the interlocutor wrongly pronounced the word *Berlin*. If this type of metalinguistic negation is what one finds in (40a), then the negation in the first sentence should be targeting the form of the sentence: it should communicate something along the lines of "the sentence giving rise to homogeneity should not be used here". In that case, one can paraphrase (40a) as in (40b).

(40) a. It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.
b. The sentence giving rise to homogeneity should not be used here. Kimberly and Helen did the dishes.

It remains unclear, however, why the (distributive vs. non-distributive) interpretation of the predicate should influence the acceptability of second sentence in (40b). In light of the above, even if the negation in our experiments was interpreted as a form of expression focus under metalinguistic negation, the results presented here are still problematic for the homogeneity approach.

21 There is another conceptual problem with Büring & Križ's (2013) formulation of homogeneity which prevents us from adopting it, namely the treatment of the assertion. As they write in their paper, Büring & Križ (2013: p. 10) predict the sentence '*It was Bill who carried the piano.*' to suffer from a presupposition failure (in Križ's (2017) terms, undefinedness) in the situation in which Bill and Fred carried the piano and neither of them did it alone. However, if the assertion of the cleft is $Bill \in [\![carry\ the\ piano]\!]$, then the assertion is false. The problem is that the sentence cannot be undefined (i.e., neither true nor false) when its assertion is false. If, on the other hand, the assertion is that $P \in \oplus [\![carry\ the\ piano]\!]$, then the assertion is true and the sentence turns out to be neither true nor false, as claimed by Büring & Križ (2013). However, if this is so then even under a metalinguistic analysis of negation in our experimental studies no difference in the acceptability of distributive vs. non-distributive predicates is expected; see Renans 2016a for discussion.

Yet another possibility is to assume that the negation in our experiments was not a truth-conditional negation that swaps truth and falsity leaving undefinedness intact, but a negation that targets the third, undefined value. In fact, one proposal in the literature in which negation targets undefinedness in a trivalent logic comes from Spector & Sudo (2017).[22] Under their account, the so-called *weak* negation maps the third undefined value # in (41) to truth in (42).

(41) $[\![\phi]\!]$ =
true iff $[\![\phi]\!] = 1$
undef iff $[\![\phi]\!] = \#$
false iff $[\![\phi]\!] = 0$

(42) Weak Negation $[\![\sim \phi]\!]$ =
true iff $[\![\phi]\!] = 0$ or $[\![\phi]\!] = \#$
false otherwise
[based on (25) in Spector & Sudo 2017]

Now consider the truth conditions of (6) according to Križ (2017), repeated in (43) for the reader.

(43) It was Kimberly who did the dishes.
true iff Kimberly is identical to the mereological sum of people who did the dishes
≈ *iff Kimberly and only Kimberly did the dishes*
false iff Kimberly does not overlap with the mereological sum of people who did the dishes
≈ *iff Kimberly did not participate in doing the dishes*
undef otherwise
≈ *iff Kimberly and somebody else did the dishes*

Under a weak negation operator, the cleft in (43) has the truth-conditions in (44).

(44) It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.
true iff Kimberly does not overlap with the mereological sum of people who the dishes, or—as in our contexts—Kimberly is properly contained in the mereological sum of people who did the dishes
≈ *iff Kimberly did not participate in doing the dishes or Kimberly and somebody else did the dishes*
false iff Kimberly is identical to the mereological sum of people who did the dishes
≈ *iff Kimberly and only Kimberly did the dishes*

Thus, whether with distributively or non-distributively interpreted predicates, (44) is true in our experimental contexts, since Kimberly is properly contained in the mere-

22 Note that Spector & Sudo's (2017) account was particular for negation in a different domain, i.e., in the exclusion of alternatives by an exhaustivity operator when an implicature arises in a presupposition.

ological sum of dish-washers including Kimberly and Helen. Therefore, again no difference in the acceptability of sentences with distributively and non-distributively interpreted predicates is predicted, contrary to what we found.

A final possibility is to assume Križ's (2016) pragmatic principle, which maps the undefined third truth-value to truth or falsity if the context makes the undefined sentence *true* or *false enough*, respectively. Under this account, however, it remains unclear why with distributively interpreted predicates *it*-clefts should be interpreted as *false enough* but with non-distributively interpreted predicates as *true enough*, if for both types of predicates the same type of analysis in terms of mereological terms is proposed.

To sum up, even by treating the negation in our experiments as a metalinguistic or weak negation targeting the third, undefined value, several challenges for the homogeneity account remain, challanges which are not faced by the alternative-based approach.

3.5 Conclusion

In this paper, we reported the results of three experiments which found that the acceptability of both *it*-clefts and their SVO counterparts in the series '*It's not* α *that did P.* α *and* β *did P.*' were reliably influenced by the distributive vs. non-distributive interpretation of the predicate. These results are consistent with the alternative-based approach to *it*-cleft exhaustivity, such as that in Velleman et al. 2012, since the acceptability of clefts—similar to SVO sentences—is predicted to differ across distributively vs. non-distributively interpreted predicates. On the other hand, the results pose challenges to the homogeneity approach, which predicts no differences for *it*-clefts across distributively and non-distributively interpreted predicates.

Chapter 4

nà-Cleft (non-)exhaustivity: Variability in Akan*

Abstract This paper presents two studies on the exhaustive inference associated with focus-background *nà*-clefts in Akan (among others, Boadi 1974, Duah 2015, Grubic, Renans & Duah 2019, Titov 2019). Despite the unforeseen response patterns in Akan for the incremental information-retrieval paradigm used in the studies, in an exploratory analysis I directly compare the results for Akan to recent experimental work on German *es*-clefts employing an identical design (De Veaugh-Geiss et al. 2018). The results are compatible with a parallel approach both (i) cross-sententially between clefts of the form 'α *nà* P' ('It is α who did *P*') and definite pseudoclefts (definite descriptions with identity statements) of the form '*Nipa no a* P *ne* α' ('The person who did *P* is α') (Boadi 1974, Ofori 2011); and (ii) cross-linguistically between Akan and German. Moreover, the by-participant variability found in both languages can be accounted for with the discourse-

* I would like to thank first and foremost Reginald Akuoko Duah and Joana Serwaa Ampofo, who not only assisted me at various stages of implementing these experiments at the University of Ghana at Legon but also showed me what a beautiful country Ghana is. I would also like to thank Ishmael Adjei, Kwaku Osei-Adjei, and Paul Okyere Omane for their assistance in preparing the materials used in the experiments, as well as Justina Lartey, Stella, and Vanessa for their help as language consultants. Moreover, I am very thankful to Agata Renans and Carla Bombi for their support and input, not to mention their company during our stay in Legon. It is not an exaggeration to say that without all of the aforementioned people this study would not have been possible. Finally, this paper has benefited by critical feedback from, in alphabetical order, Carla Bombi, Reginald Akuoko Duah, Mira Grubic, Mareike Philipp, Agata Renans, Laine Stranahan, Marta Wierzba, and Malte Zimmermann. This research was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 317633480—SFB 1287, Project C02, as well as the "Exhaustivity in *it*-Clefts" project in the XPrag.de priority program (SPP 1727). The usual disclaimer applies: any and all errors are entirely mine.

pragmatic analysis of cleft exhaustivity in De Veaugh-Geiss et al. 2018 (based on Pollard & Yasavul 2016), which has been adopted for focus-background clefts in various languages.

Keywords: Akan; *nà*-clefts; definite pseudocleft; exhaustivity; variability; experimental study

4.1 Introduction

Across numerous languages, bipartitional focus-background clefts of a form comparable to '*It is* $[\alpha]_{focus}$ [*who did* P]$_{background}$' have been argued to give rise to at least three layers of meaning: (i) the truth-conditional assertion corresponding to the canonical form 'α *did* P', (ii) the existence presupposition '*someone did* P', and (iii) an exhaustivity inference '*nobody other than* α *did* P'. Focus-marking *nà*-cleft structures in Akan are no different (among others, Ellis & Boadi 1969, Boadi 1974, Ofori 2011, Ameka 2010, Amfo 2010, Duah 2015, Pfeil, Genzel & Kügler 2015, Grubic, Renans & Duah 2019, Titov 2019). Consider example (1); of particular importance here is the exhaustivity component in (iii), which remains the focal point of an ongoing debate.[1]

(1) a. Kodwo na ɔ-di-i adua no.
Kodwo FM 3.SG.SBJ.ANI-eat-PST beans DET

1 All examples are reproduced as close as possible to the original, although I will take the liberty to make minor changes to the glosses to follow the Leipzig Glossing Rules (Max Planck Institute for Evolutionary Anthropology 2018) or to be consistent, but only when doing so does not appear to be critical. All examples include a citation of the primary source and the example number there to assist the reader in consulting the original directly. Note that occasionally an Akan native-speaker consultant disagreed with the orthographic form used in the example sentences, and any changes I have made will be marked in square brackets when the Akan text presented here has been modified. Although high and low tones are not marked in the standard orthography of Akan (Kobele & Torrence 2006), tones are marked in the examples when the cited text or language consultant included them. The following glosses will be used: 1/2/3 = first/second/third person, ANI/INA = animate/inanimate, DET = determiner, FM = focus marker, NEG = negative, PRT = particle, PRS/PST = present or habitual/past or completive (with the use of this glossing I wish to remain neutral with respect to aspectual reference), SBJ/OBJ = subject/object, SG/PL = singular/plural. I will follow standard practice and use * and # to mark syntactic and semantic/pragmatic unacceptability, respectively, and ? to mark uncertainty or questionability. Moreover, for a unified presentation I will occasionally modify the acceptability judgments from the cited source to be consistent (e.g., changing * from the original to # here when the reported unacceptability is due to semantics/pragmatics and not syntactic ill-formedness).

'*It was Kodwo who ate the beans.*'
[adapted from ex. (6) in Pfeil, Genzel & Kügler 2015]

(i) Kodwo ate the beans. (canonical: *asserted*)
(ii) Someone ate the beans. (existential: *presupposed*)
(iii) Yaa, Kofi, Abena … did not eat the beans, i.e., nobody other than Kodwo ate the beans. (exhaustivity: *?*)

Grubic, Renans & Duah (2019), Duah (2015), Pfeil, Genzel & Kügler (2015), Ameka (2010), among others, have all claimed that when focus-marked entities in Akan appear *ex situ*—that is, when the structure in (1) is involved—the predicate is interpreted as only holding for the focused constituent appearing in the left periphery, the exhaustive reading.[2] As Titov (2019: p. 15) phrased it: for a true sentence "[the] exhaustive reading results in an interpretation according to which propositions containing an alternative to the focus receive an opposite truth-value", illustrated in (iii) above.

Although there is broad agreement that focus-background clefts give rise to such an exhaustive meaning cross-linguistically, there remains no consensus regarding the nature of this inference: Is exhaustivity conventionally-coded as part of the cleft-structure, with the exhaustive inference typically taken to be a presupposition in current literature (e.g., Duah 2015, Grubic, Renans & Duah 2019 for Akan, as well as the theoretical accounts in Velleman et al. 2012, Büring & Križ 2013)? Or is exhaustivity a non-conventionally-coded layer of meaning such as an implicature (Horn 1981, 2014, De Veaugh-Geiss et al. 2015; see also Titov 2019) or a discourse-pragmatic phenomenon (Pollard & Yasavul 2016, De Veaugh-Geiss et al. 2018; see De Veaugh-Geiss et al. 2018 for an overview of the various semantic/pragmatic approaches)?

Against this background, several recent studies have compared experimentally the exhaustive inference of clefts to exhaustive inferences in other constructions such as sentences with exclusives, *in situ* focus constructions with canonical non-cleft word order, or definite pseudoclefts (definite descriptions with identity statements of the form '*The one who did* P *is* α') (e.g., Destruel et al. 2015, Onea & Beaver 2009, Xue & Onea 2011, Byram-Washburn, Kaiser & Zubizarreta 2013, De Veaugh-Geiss et al. 2015, Destruel & De Veaugh-Geiss 2019). Although the debate regarding

2 Also referred to in the literature as *exhaustiveness* (e.g., Horn 1981, Declerck 1988, Drenhaus, Zimmermann & Vasishth 2011, Patten 2012), *exhaustivity* (e.g., Schulz & Van Rooij 2006, Büring & Križ 2013), or *exhaustive listing* (Horn 1981), as well as *uniqueness* (Delin & Oberlander 2005) and *exclusiveness* (Collins 1991). For Akan specifically, the terms *exhaustive focus marking* (Duah 2015), *exhaustive focus* (Titov 2019), *selective focus* (Amfo 2010; in the sense of Dik et al. 1981), and *exclusive focus* are also found. This list, in part from De Cesare & Garassino 2015, is certainly not exhaustive.

the nature of cleft exhaustivity is far from settled, the results of the experimental literature are generally consistent with the following claims: (i) exhaustivity in clefts is weaker than in sentences with exclusives but stronger than in canonical non-cleft constructions; (ii) cleft exhaustivity need not hold in all contexts and thus it is not conventionally-coded as part of the meaning of the cleft structure; and finally, (iii) clefts and definite pseudoclefts generally exhibit parallel behavior when compared directly (although see De Veaugh-Geiss et al. 2015). Furthermore, and of interest here, the above claims appear to hold cross-linguistically (e.g., German, English, and to some extent French; but see Destruel & De Veaugh-Geiss 2019). The experimental studies here—which employ the same design as the truth-value judgment task in violation contexts reported for German (De Veaugh-Geiss et al. 2018)—extend this work to *nà*-clefts in Akan.

Despite the unforeseen response patterns in the incremental information-retrieval paradigm used in the experiments, an exploratory analysis of the results suggests striking parallels between Akan and German. Although several variables were recorded in the experiments—including the proportion of continuations in the stepwise presentation of information as well as final truth-value judgments—in this manuscript I only analyze the latter since it is the final truth-value judgments which were ultimately informative (cf. the analysis in De Veaugh-Geiss et al. 2018). Anticipating the discussion of the results, although the analysis of the final truth-value judgments shows intermediate exhaustivity patterns for clefts and definite pseudoclefts on avarage in both languages, a by-participant analysis reveals a bimodal distribution of responses, supporting the view that exhaustivity is an inference that consistently arises for some speakers and consistently does not arise for others.

This paper will proceed as follows. Section 4.2 presents the background on Akan as well as the four sentence types of interest: *nà*-clefts, canonical Subject-Verb-Object (SVO) sentences, exclusives, and definite pseudoclefts. Following this, Section 4.3 describes the experiments testing *na*-cleft and definite pseudocleft exhaustivity, with exclusives and non-cleft SVO sentences as controls; moreover, I present a post hoc exploratory analysis of the results, with a direct comparison to the results of the studies on clefts and pseudoclefts in German reported in De Veaugh-Geiss et al. 2018. In short, no difference was found between either clefts and pseudoclefts, nor German and Akan. A discussion of these results is in Section 4.4, in which—following the analysis for German in De Veaugh-Geiss et al. 2018 (cf. Pollard & Yasavul 2016)—the discourse-pragmatic approach is proposed to account for the data. Section 4.5 concludes.

4.2 Background

4.2.1 The Akan language

With over 8 million speakers, Akan—a cover term for a group of mutually-intelligible dialects—is the most widely spoken language in central and southern parts of Ghana (Lewis, Simons & Fennig 2013, Duah 2013, Bombi 2017). The language belongs to the Kwa subfamily in the Niger-Congo phylum and includes the dialects Agona, Akuapem Twi, Akyem, Asante Twi, Brong, Fante, Gomua, Kwahu, and Wasa (Dolphyne & Kropp Dakubu 1988: p. 52).[3] The three largest dialects include Asante Twi (3.8 million speakers), Fante (2.7 million speakers), and Akuapem Twi (600,000 speakers) (Lewis, Simons & Fennig 2013).

Written Akan has a standardized orthography and is taught in schools, but since the standard is not rigorously enforced there are dialectal differences found in the written form (Duah 2013: pp. 5–6; cf. discussion in Bombi & De Veaugh-Geiss 2018). And although English is the official language of Ghana, English is used mostly by educated speakers when there is no other common language; by contrast, Akan is widely used by both literate and non-literate people (Duah 2013: p. 6). In fact, in many areas, such as Greater Accra, Ashanti, Eastern, Brong Ahafo, Western and Central regions of Ghana, Akan is the dominant language in broadcasting (Duah 2013: p. 7). As Osam (2004) claims: "even though no official declaration has been made, Akan is growing in its influence as a potential national language, especially since people who speak other languages sometimes use it as a *lingua franca*."

Like many of the languages in the Kwa subfamily, Akan is a tone language with a high and a low tone—marked with an acute and a grave accent, respectively—and these tones have lexical and grammatical functions (Dolphyne 1988, Dolphyne & Kropp Dakubu 1988, Genzel 2013). The Akan language displays head-initial characteristics, such that nouns come before determiners, adjectives, and numerals (Kobele & Torrence 2006, Boadi 2005, Aboh 2010), as illustrated in (2).

(2) Kontromfi no tua dua kakraa futufutu.
monkey DET possess.PRS tail huge fluffy
'*The monkey has a huge fluffy tail.*' [ex. (1) in Genzel 2013]

4.2.2 Two focus-marking strategies in Akan: *in situ* vs. *ex situ*

In this section I will first introduce the notion of focus. This will be followed by a presentation of two focus-marking strategies in Akan, namely, the *in situ* strategy and

3 See Cahill 1985, Dolphyne 1988, and Abakah 2005 for details regarding differences between the dialects, specifically at the segmental and tonal level (Genzel 2013).

ex situ strategy with the focus marker *nà*. Then—following Boadi (1974), Grubic, Renans & Duah (2019) (inter alia), who take the *ex situ* strategy with *nà* to be a cleft construction—I will describe the *nà*-cleft in terms of its structure. Finally, I will present the claims of a subject/non-subject asymmetry for focus marking in Akan.

4.2.2.1 Focus

Focus is an information-structural notion that, following Rooth (1992), I take to indicate the presence of salient alternatives in context relevant for interpretation (see, e.g., Krifka 2008). For example, one way in which alternatives can be relevant for interpretation is in question-answer pairs, i.e., new-information focus, as in (3a); other ways include corrective/contrastive focus, as in (3b), and selective focus, as in (3c), among others (Hartmann & Zimmermann 2007: p. 366). Note that the focus is indicated by the subscript $_F$.

(3) a. Who was liberated yesterday? [Simona]$_F$ was liberated yesterday.
b. Peter bought a Mercedes. No, he bought a [Toyota]$_F$.
c. Did you have bagels or muffins for breakfast? I had [bagels]$_F$ for breakfast.

[ex. (1a)–(1c) in Hartmann & Zimmermann 2007]

In the following discussion of focus-marking in Akan, I will concentrate primarily on new-information focus; see, e.g., Genzel 2013, Grubic, Renans & Duah 2019, Titov 2019 for additional uses of focus constructions in Akan, in particular corrective/contrastive focus.

4.2.2.2 Focus-marking in Akan

Akan employs at least two strategies for marking focus. First, there is an *in situ* strategy, in which the focus-marked constituent (or simply, the focus) appears in base position in the canonical SVO sentence, illustrated in the answer A in (4).

(4) Q: Deɛbɛ[n] na Kofi di-i?
what FM Kofi eat-PST
'*What did Kofi eat?*'
A: Kofi di-i [adua no]$_F$.
Kofi eat-PST beans DET
'*Kofi ate the beans.*'

[modified from ex. (35) in Genzel 2013]

In such cases, Ofori (2011: p. 254) writes that there is "prominence (loudness) that picks the focused unit whenever [...] overt restricting mechanisms (especially, the

[*ex situ* strategy, as below]) are not being used"; cf. Kügler & Genzel (2012: p. 727), who claim that "Akan shows a tendency to lower the intensity of the post-focal part, which has been interpreted as indirect way to make the focused element prominent or to align it" (see also Marfo & Bodomo 2005 and Titov 2019).

Second, there is an *ex situ* strategy, in which focus is marked morphosyntactically by means of the focus marker *nà* and the focus immediately precedes it in the left periphery,[4] illustrated in the answer A in (5) (among others, Kobele & Torrence 2006, Ameka 2010, Grubic, Renans & Duah 2019, Titov 2019).

(5) Q: Hwáń nà Nti bó-ò nó?
who FM Nti beat-PST 3.SG.OBJ
'*Who did Nti beat?*'
A: [Kwaku$_i$]$_F$ nà Nti bó-ò nó$_i$.
Kwaku FM Nti beat-PST 3.SG.OBJ
'*It was Kwaku that Nti beat.*' [ex. (2) in Duah 2015]

Note that when the focus-marked word or phrase appears at the left edge of the clause, as in answer A in (5), I will follow previous literature by referring to an *ex situ* strategy, although I wish to remain neutral as to whether the constituent is moved from the canonical sentence or base-generated in the left periphery.[5]

The question-answer pairs in (4)–(5) are examples of narrow focus, in which the new information answering the *wh*-question is a smaller part of the larger sentence it appears in; in some cases, though, the entire sentence may include new information. This is the case, for instance, in many out-of-the-blue contexts or in answers to questions such as *What happened?*, referred to as wide or broad focus (Selkirk 1984). Broad focus in Akan is expressed by use of the canonical SVO sentence, as illustrated by the acceptability of A1 as an answer to the question in (6); by contrast, in answer to the same question, the *ex situ* construction A2 is unacceptable.[6]

4 In this paper I will ignore the morphosyntactically-comparable marker *de(ɛ)*, which also precedes the element in the left periphery and is sometimes analyzed as a focus marker in the literature (e.g., Boadi 1974; however, see Titov 2019 for claims that this element rather marks contrastive topic). Regarding the focus marker *nà*, Duah (2015: p. 2) notes, "*nà* always occurs with a low tone (`) and can be distinguished from the clausal conjunction *nà* and the past discourse marker [*ná*]" (see also Amfo 2007). In a recent analysis, Duah (2019: p. 9) proposes that the three uses of *na* particles in Akan in fact "have a single underlying discourse particle [. . .] as its source".

5 See Saah 1994, 2010, Ofori 2011 for non-movement accounts, and Titov 2019 for additional tests supporting this position; cf. Renans 2016a for discussion of base-generation of the focus in the left periphery for Ga, a Kwa language also spoken in Ghana.

6 Cf. broad focus with French subject clefts, as discussed by Destruel & De Veaugh-Geiss (2018: §2.1) (see also Katz 2000, Lambrecht 2001). It may, however, be possible to get wide focus *ex situ* in Akan with perfect aspect (used to narrate events). I thank Reginald Akuoko Duah for bringing this to my attention.

(6) Q: Deɛn na e-si-i?
what FM 3SG.INA-happen-PST
'*What happened?*'
A1: Adwoa dii aduan no.
Adwoa ate food DET
'*Adwoa ate the food.*'
A2: #Aduan no na Adwoa dii.
food DET FM Adwoa ate
'*It is the food Adwoa ate.*' [ex. (23) in Titov 2019]

Titov (2019: p. 11) writes that the difference in acceptability between A1 and A2 in (6) "suggest[s] that focus projection is available only from the in-situ object position but never from the leftmost object position in Akan". In other words, the focus-marking strategy with *nà* is unacceptable for A2 since the information sought by the questions must be expressed by a constituent larger than the focus-marked one, but focus cannot project from the *ex situ* position.[7]

4.2.2.3 Structure of the *ex situ* (cleft) construction

The sentence with *nà* partitions the clause into two parts: the focus-marked *new* information preceding the focus marker, followed by the *given* information in the predicate (e.g., Grubic, Renans & Duah 2019; see also Zimmermann et al. to appear). In the spirit of Boadi (1974), Kobele & Torrence (2006), Grubic, Renans & Duah (2019), among others, I will refer to these *nà*-constructions as clefts, a schematic representation of which is shown in (7).

(7) $[_{\text{cleft pivot}}\ \alpha]$ nà $[_{\text{cleft predicate}}$ P] (no) (*nà*-cleft)

Some terminology regarding the *nà*-cleft structure would be helpful at this point. The constituent α preceding the focus marker is the *cleft pivot* (or simply *pivot*) and the clause *P* following it the *cleft predicate*. Except for negated clefts (which is not relevant for the studies here), the occurence of a copula ɛyɛ 'it is' is optional in *nà*-clefts,[8] in contrast to English *it*-clefts or German *es*-clefts, among other

7 De Veaugh-Geiss et al. (2015: §4) derive the comparatively strong exhaustive implicature in *it*-clefts from this lack of focus projection; see Križ 2017: §2.1 for arguments against this approach.

8 Boadi (1974: p. 14) writes that the copula ɛyɛ "has no semantic content", and Duah (2015: p. 19) claims that "there is no difference between [(i) below without the copula] and [with the copula] in terms of exhaustivity of focus, but [(i) without the copula] has become the default construction for ex situ focus with *nà*"; however, see Ofori 2011: §4.1 for arguments against such claims.

(i) (ɛ-yɛ) Kofi nà ɔ-bá-à há.
3.SG.SBJ.INA-be Kofi FM 3.SG.SBJ-come-PST here
'*It was Kofi who came here.*' [based on ex. (25) in Duah 2015]

languages. Moreover, an optional determiner-like element, *no*, may appear in clause-final position of the cleft predicate. The contribution of this element, which is homophonous to the definite determiner (Kobele & Torrence 2006; see also Arkoh & Matthewson 2013: §4.1),[9] is debated. Kobele & Torrence (2006: Fn. 7) comment that it "adds some type of 'emphasis'", while Boadi (1974) writes that "the speaker intends to remind the hearer that the incident has been referred to earlier". See discussion of this clause-final determiner-like element in Bombi et al. 2019 (for Akan) and Grubic, Renans & Duah 2019: §8.1 (for a comparison of Akan and Ga).

When the pivot is animate, a resumptive pronoun, coindexed with the *ex situ* focus, appears in base position in the cleft predicate. This is illustrated in A in (8) by the third person singular resumptive pronoun *ò-* prefixed to the verb and coindexed with the subject in the left periphery. (For an example of the resumptive pronoun with *ex situ* object focus, see A in (5) on page 133, i.e., the third person singular object pronoun *nó*.)

(8) Q: Hwáń nà ò-dí-ì àdùàné nó?
who FM 3.SG.SBJ-eat-PST food DET
'*Who ate the food?*'
A: Kofi$_i$ nà ò$_i$-dí-ìɛ́.
Kofi FM 3.SG.SBJ-eat-PST
'*It was* KOFI *who ate it.*' [ex. (2) in Duah 2015]

The resumptive pronoun is absent, however, when the constituent is inanimate, and including one may result in ungrammaticality (Fiedler & Schwarz 2007, Duah 2015), shown in (A) in (9).

(9) Q: Dɛ́ń nà Kofi dí-ìɛ̀?
what FM Kofi eat-PST
'*What did Kofi eat?*'
A: ɛ̀mó nà Kofi dí-ìɛ̀ (*nó).
rice FM Kofi eat-PST 3.SG.OBJ
'*It was* RICE *that Kofi ate.*'
[adapted from ex. (4) in Duah 2015]

---

9 When the optional clausal *no* and the resumptive pronoun *no* appear in the same clause, the result is the doubling of *no*.

(i) Onipa$_i$ a mehuu no$_i$ no.
'*The person that I saw . . .*' [Footnote 3 in Ofori 2011]

Thus far, all example sentences have been for subject and object focus, but *ex situ* clefting as a focus-marking strategy is available for other constituents as well, such as verbs, as in (10), realized with verb copying, and adjuncts, as in (11).

(10) Q: Dɛ́ń nà Yaw tɛ́-ɛ̀ Akwasi?
what FM Yaw do-PST Akwasi
'*What did Yaw do to Akwasi?*'
A: Píá nà Yaw píá-à Akwasi.
push FM Yaw push-PST Akwasi
'*It was* PUSHING *that Yaw did to Akwasi.*'
[ex. (3) in Duah 2015]

(11) Q: Dàbɛ́ń nà Ama bá-àɛ́?
day.which FM Ama arrive-PST
'*When did Ama arrive?*'
A: ɛ̀nórà nà Ama bá-àɛ́.
yesterday FM Ama come-PST
'*It was* YESTERDAY *that Ama came.*' [ex. (5) in Duah 2015]

4.2.2.4 Subject/non-subject asymmetry

For non-subjects, the focus may appear either *in situ* in the SVO sentence, as in (12), or *ex situ* in a *nà*-cleft, as in (13).

(12) *SVO* (*in situ*)
Q: '*What did Kofi eat?*'
A: Kofi di-i a-dua no.
Kofi eat-PST NOM-beans DET
'*Kofi ate the beans.*' [ex. (4) above]

(13) nà-*Cleft* (*ex situ*)
Q: '*What did Kofi eat?*'
A: ɛ̀mó nà Kofi dí-ìɛ̀.
rice FM Kofi eat-PST
'*It was* RICE *that Kofi ate.*' [adapted from ex. (9) above]

The same holds for non-subject interrogative sentences. In (14)–(15) the acceptability of both strategies for the interrogative *hena* 'who' is illustrated.

(14) *SVO* (*in situ*)
Q: Kofi bɔɔ hena?
Kofi hit.PST who
'*Who did Kofi hit?*' [ex. (1b) in Kobele & Torrence 2006]

(15) nà-*Cleft* (*ex situ*)

Q: Hena nà Kofi bɔɔ (no)?
who FM Kofi hit.PST 3.SG
'*Who is it that Kofi hit?*' [ex. (1c) in Kobele & Torrence 2006]

In contrast to non-subject interrogatives, only the *ex situ* strategy is available for subject interrogatives: that is, it is necessary for the *wh*-word to appear in the left periphery preceding the focus marker *nà*, as in example (17). When the *wh*-subject occurs without *nà*, the result is ungrammaticality, shown in (16) (Duah 2015: pp. 6–7; see also Titov 2019: p. 20, who describes this asymmetry for interrogatives as "absolutely uncontroversial").[10]

(16) *SVO* (*in situ*)

Q: *Hwáń pìá-à Ama?
who push-PST Ama
(intended) '*Who pushed Ama?*' [ex. (8b) in Duah 2015]

(17) nà-*Cleft* (*ex situ*)

Q: Hwáń nà ò-píá-à Ama?
who FM 3.SG.SBJ-push-PST Ama
'*Who pushed Ama?*' [ex. (8a) in Duah 2015]

Similarly, in declarative sentences focused subject constituents typically appear in the left periphery, in particular in cases of corrective/contrastive focus (Marfo & Bodomo 2005, Grubic, Renans & Duah 2019) and narrow information focus, as in (20) in answer to (18) (Duah 2015, Titov 2019); cf. the unacceptability of the *in situ* strategy in (19) in answer to the same question. In fact, given the unacceptability of answers as in (16) and (19), it has been argued that Akan displays a subject/non-subject asymmetry.

(18) Q: Hwáń nà ɔ̀-bá-à há?
who FM 3.SG.SBJ-come-PST here
'*Who came here?*' [ex. (10) (Q) in Duah 2015]

(19) *SVO* (*in situ*)

A: #Kwabena bá-à há
Kwabena come-PST here

10 In the Fante dialect, however, *wh*-subjects can appear at the left edge without a focus marker.

(i) Hwáná píá-à Ama?
who push-PST Ama
'*Who pushed Ama?*' [ex. (8c) in Duah 2015]

'*Kwabena came here.*' [ex. (10) (A1) in Duah 2015]

(20) nà-*Cleft* (*ex situ*)

A: Kwabena nà ɔ-bá-à há.
Kwabena FM 3.SG.SBJ-come-PST here
'*It was Kwabena who came here.*' [ex. (10) (A) in Duah 2015]

In short, the asymmetry claim is that focused subjects are obligatorily in the left periphery accompanied by the focus marker *nà*, whereas focused non-subjects may appear either *in situ* or *ex situ* (Fiedler et al. 2010, Fiedler & Schwarz 2005, 2007, Marfo & Bodomo 2005, Titov 2019; see also Zerbian 2007).[11] However, Genzel (2013) and Duah (2015) contradict claims of a subject/non-subject asymmetry for Akan, arguing that no asymmetry exists, "at least not in the original sense of the term", in the words of Genzel (2013: p. 207). In fact, it appears that a relevant factor for whether subject focus is realized *in situ* vs. *ex situ* is the (non-)exhaustive interpretation, the topic of the following section.

4.2.3 Exhaustivity inferences across sentence types

In this section, I will first show that both subject and object focus can be realized *in situ* with a non-cleft SVO sentence and *ex situ* with a *nà*-cleft, but with a difference in exhaustive interpretations. Following this I will discuss the asserted exhaustivity inference in sentences with the exclusive particle *ńkóáá* 'only'.

4.2.3.1 (Non-)exhaustive SVO sentences vs. exhaustive *nà*-clefts

Subject focus Duah (2015) unpacks the discourse conditions under which subject focus is claimed to be realized *in situ* in the non-cleft SVO sentence as opposed to *ex situ* with a *nà*-cleft, and concludes that exhaustive interpretations play a critical role. Consider the following sentences in context. Duah (2015: p. 21) argues that the *in situ* subject focus A in (21) is acceptable because "the relevant individuals for an answer cannot be exhaustively listed"; by contrast, the clefted subject focus in A1 is unacceptable.

11 Comparable subject/non-subject asymmetries occur in numerous African languages in various cleft and cleft-like sentences as well as other focus-marking structures (e.g., Ameka 2010, Fiedler et al. 2010). These include Niger-Congo languages (*cleft/cleft-like structures*: Byali, (Chi)Chewa, (Chi)Tumbuka, Dzamba, Northern Sotho, Yorùbá, Zulu; *focus marking*: Banda-Linda, Dagbani, Ewe, Fon), Afro-Asiatic languages (*cleft/cleft-like structures*: Afaan Oromoo, Bura, Goemai; *focus marking*: Bole, Hdi), and Nilo-Saharan languages (cleft/cleft-like structures: Lango; *focus marking*: Koyraboro Senni, Koyra Chiini, Masalit). This partial list was compiled when working as an undergraduate research assistant to Gisbert Fanselow in the SFB 632 'Information Structure'; for a written summary, please contact the author.

(21) Q: Hwáń nà ɔ̀-bá-à àyíé nó?
who FM 3.SG.SBJ-come-PST funeral DET
'*Who attended the funeral?*'
A: *SVO* (*subject focus in situ*)
Kofi bá-àɛ́.
Kofi come-PST
'*Kofi came.*'
A1: nà-*Cleft* (*subject focus ex situ*)
#Kofi nà ɔ̀-bà-àɛ́.
Kofi FM 3.SG.SBJ-come-PST
'*It was Kofi who came.*' [ex. (27) in Duah 2015]

Duah (2015) employs various diagnostics to test for exhaustivity, including co-ordination tests, *mention-some* answers, co-occurance with additive particles, and corrections with '*No, also . . .* ' (see Duah 2015: §3 for details; cf. Fominyam & Šimík 2017: §4.5 for related diagnostics applied to Awing, a Grassfields Bantu language). Based on such examples, Duah (2015: p. 21) concludes that "subject in situ focus is possible in some focus contexts[; however] when the context does not require that what is focused be *non-exhaustive*, focus tends to be ex situ for subjects" (emphasis added).[12]

Subject focus with the *nà*-cleft is argued to give rise to a robust exhaustive inference.[13] This is illustrated in (22), in which the exhaustive inference from the first sentence—i.e., Nti and no one else went to school—is shown to be difficult to cancel by the second sentence (see, e.g., Saur 2013 and De Veaugh-Geiss et al. 2015 for experimental studies using similar tests with German *es*-clefts).

(22) Q: Who went to school yesterday?

12 The claims in Duah 2015 are compatible with the results of the pre-study questionnaire and production task reported in Pfeil, Genzel & Kügler 2015. Although the number of language consultants in the pre-study was small, there was consensus: all three translated answers to non-exhaustive questions (e.g., *Who else arrived?*) with *in situ* subject focus (see, however, Titov 2019: §4 for a different interpretation of these results). By comparison, when the question elicited an answer that is intended to be exhaustive (e.g., *Who ate the food?*), two speakers commented that "the in situ construction [. . .] did not answer the question [. . .] sufficiently" and "one speaker described [*in situ* subject focus] as odd" (Pfeil, Genzel & Kügler 2015: pp. 92–93); instead, the clefting strategy was preferred. Furthermore, the results of the production task reported there show a similar pattern: in the non-exhaustive condition, participants produced subject focus *in situ* a majority of the time, whereas in the exhaustive condition participants produced subject focus *ex situ* a majority of the time. Thus, Pfeil, Genzel & Kügler (2015: p. 104) conclude that "the results of the production experiment clearly indicate that the choice of the focus marker [*nà*] depends on exhaustivity".

13 Specifically, clefts in declarative sentences; see Duah 2015: pp. 22–25 for discussion of exhaustivity, and the lack thereof, in clefted interrogatives in Akan.

A: #Nti nà ɔ̀-kɔ́-ɔ̀ sùkúù ɛ̀nórà. ɛ́nà Yaw ńsó kɔ́-ɔ́
Nti FM 3.SG.SBJ-go-PST school yesterday and Yaw also go-PST
sùkúù.
school
'*It was Nti who went to school. And Yaw also went to school.*'
[ex. (47) in Grubic, Renans & Duah 2019]

Object focus By contrast, focused objects *in situ* with canonical SVO word order may give rise to an exhaustive interpretation in contexts that enforce a narrow information focus, according to Grubic, Renans & Duah (2019: pp. 15–16). For instance, since it is known in the context in (23) that Abena ate more than just corn, the answer in (23) is unacceptable.

(23) *Context*: Abena ate corn and groundnuts.
Q: What did Abena eat?
A: *SVO* (*object focus in situ*)
#Abena dì-ì àbùró.
Abena eat-PST corn
'*Abena ate corn.*' [ex. (35) in Grubic, Renans & Duah 2019]

Nevertheless, the exhaustive reading for *in situ* focus is argued to be a non-conventionally-coded, pragmatic inference; therefore, it is cancellable and as a result the series of sentences in (24) is acceptable (Grubic, Renans & Duah 2019).

(24) Q: What did you buy?
A: Mè-tɔ̀-ɔ̀ àtààdéɛ́ ɛ́nà mé-tɔ́-ɔ̀ɛ́ m̀pàbòá ńsó.
1.SG-buy-PST short and 1.SG-buy-PST shoes also
'*I bought a shirt and I bought shoes also.*'
[ex. (44-A2) in Grubic, Renans & Duah 2019]

In comparison to *in situ* object focus, *ex situ* object focus with a *nà*-cleft gives rise to a robust exhaustive inference which is argued not to be cancellable, illustrated in example (25). This strong exhaustive effect is directly comparable to that for subject focus shown in (22) above.

(25) Q: What did you buy?
A: #Àtààdéɛ́ nà mè-tɔ́-ɔ̀ɛ́ ɛ́nà mè-tɔ̀-ɔ̀ m̀pàbòá ńsó.
short FM 1.SG-buy-PST and 1.SG-buy-PST shoes also
'*It was a shirt that I bought and I bought shoes also.*'
[ex. (44-A1) in Grubic, Renans & Duah 2019]

Importantly, in comparison to subject focus—for which the *nà*-cleft is the default strategy unless the context requires a non-exhaustive reading—for object focus the *in situ* focus-marking strategy has been claimed to be the default, with the *ex situ* strategy associated with contrastive focus (Genzel 2013) or strong exhaustivity (Duah 2015). Genzel (2013: p. 208) writes: "[in Akan] focused objects are frequently realized in their base position, i.e. without any additional syntactic and morphological marking[; thus,] the in-situ strategy is the preferred one for focused objects." I will return to this point in the hypotheses presented in Section 4.2.5.

Do *nà*-clefts conventionally-code exhaustivity? Whereas Duah (2015: p. 25) argues, following Boadi (1974), that "the focus particle *nà* can be appropriately identified as an *exhaustive focus particle* because it occurs only in exhaustive focus environments", Pfeil, Genzel & Kügler (2015) contradict such a strong claim, since in at least some cases in their experiments *ex situ* focus marking occurred in non-exhaustive contexts as well (albeit to a much lesser degree). Pfeil, Genzel & Kügler (2015: p. 104) write: "our data suggest that this interpretation of the focus marker [i.e., as an exhaustive focus marker] is too narrow, since the use of [*nà*] is possible in non-exhaustive contexts as well".

In a similar vein, Titov (2019) contends that *nà*-clefts do not truth-conditionally encode exhaustivity. As argued in Titov 2019, for instance, the second clause in example (26) with the concessive *mmom*—indicating "that the first clause denotes a circumstance which might be expected to preclude the action of the second clause, but does not" (p. 16)—would be semantically odd if the exhaustivity inference was encoded as part of the truth-conditions of the *nà*-cleft structure.[14]

(26) ɔbaa no na me-huu no na mmom m-a-n-hu obi
woman the FM I-saw 3.SG but conversely I-not-saw [sic] person
fo[f]oro biara.
new any
'*Although the woman is who I saw, I didn't see anyone else (any new person).*"

[ex. (31) in Titov 2019]

In short, based on (26) among other examples, Titov (2019: p. 16) claims that "the exhaustive reading is not part of the truth-conditional interpretation of the focus construction [with *nà*-clefts] and is therefore not obligatory".

14 Reginald Akuoko Duah (p.c.) points out the sentence with the exclusive corresponding to (26) (i.e., ɔbaa no nkoaa na me-huu no na mmom . . . , '*I only saw the woman, but conversely* . . . ') behaves the same way, which would be problematic for the argumentation in Titov 2019.

Grubic, Renans & Duah (2019: §5.2.2) conclude, however, that the exhaustive inference must be conventionally-coded as part of the meaning of the *nà*-cleft structure since the inference is not cancellable, as shown in (22) and (25). Thus, they model the exhaustive inference as a presupposition in their analysis. A presuppositional approach to *nà*-cleft exhaustivity follows a range of analyses found in the theoretical literature for *it*-clefts, including, but not limited to, Percus 1997, Velleman et al. 2012, Büring 2011, Büring & Križ 2013. I return to these approaches, in particular the semantic definite accounts of Percus 1997, Büring & Križ 2013, etc., shortly. First, however, I will discuss the asserted exhaustivity for sentences with the exclusive *ńkóáá* 'only' in the following section.

4.2.3.2 Exclusive focus particle: asserted exhaustivity

Using the exclusive particle *ńkóáá* 'only' is another way to express an exhaustive answer in Akan. Consider the example in sentence (27), in which the exclusive appears with the focus-marked constituent in the left periphery.[15]

(27) Kwame ńkóáá nà ɔ-kɔ́-ɔ fíé.
Kwame only FM 3.SG.SBJ-go-PST home
'*Only Kwame went home.*' [ex. (ia) in Footnote 7 in Duah 2015]

As with exclusive particles in other languages, *ńkóáá* 'only' is claimed to assert exhaustivity. One diagnostic for teasing apart assertions from presuppositions or pragmatic implicatures is the so-called *reason-clause* test from Beaver & Clark (2008). As discussed in Grubic, Renans & Duah 2019, only asserted information in a *because*-clause can be understood to be the reason for the main clause. Thus, for a *because*-clause with *ńkóáá* 'only', as in (28), the reason for repeating the exam is the asserted exhaustivity reading, i.e., "the teacher doesn't want other students to fail" (Grubic, Renans & Duah 2019: p. 241).

(28) *Context*: The teacher will repeat the exam, . . .
a. èfisɛ́ Yaw ńkóáá nà ɔ-twá-à ǹsɔ́hwɛ́ nó.
because Yaw only FM 3.SG.SBJ-pass-PST exam DET

15 It is worth pointing out here—as Duah (2015: Fn. 7) notes—that the exhaustive inference associated with the exclusive particle is independent of the *ex situ* focus marking with *nà*. This is illustrated by the exhaustive interpretation of the *in situ* use of the exclusive in (i).

(i) Arko tɔ̀-ɔ àsɔ́màdéɛ́ ńkóáá.
Arko buy-PST earrings only
'*Arko bought only earrings.*' [ex. (23a) in Duah 2015]

Anticipating the discussion of materials in Section 4.3.1.2, the *ex situ* strategy with the exclusive particle *ńkóáá* was used for the stimuli used in the experiments.

'*because only Yaw passed the exam*'
(Comment: the teacher wants everybody to pass, and since Yaw is the only one who passed, the teacher will repeat the exam)
[ex. (41a) in Grubic, Renans & Duah 2019]

By contrast, for a *because*-clause with a *nà*-cleft, as in (29), the reason for repeating the exam is not the exhaustive inference, but rather the canonical meaning, i.e., that Yaw passed the exam.

(29) *Context*: The teacher will repeat the exam, . . .

a. èfisɛ́ Yaw nà ɔ̀-twá-à ǹsɔ́hwɛ́ nó.
because Yaw FM 3.SG.SBJ-pass-PST exam DET
'*because it was Yaw who passed the exam*'
(Comment: the teacher doesn't want Yaw to pass the exam, so he will repeat the exam)
[ex. (41b) in Grubic, Renans & Duah 2019]

Additionally, Grubic, Renans & Duah (2019) report that the exhaustive inference is visible to negation for sentences with the exclusive *ńkóáá* 'only', but not for *nà*-clefts (see §5.2.1 in Grubic, Renans & Duah 2019 for details). Thus, given the above diagnostics, the authors conclude that exhaustivity is asserted for exclusives but not for clefts in Akan.

4.2.4 Definite pseudoclefts (and clefts)

In this section I discuss definite pseudoclefts, which have been argued in the literature to share the same underlying syntax and semantics of clefts. I will begin with an overview of the discussion for definite pseudoclefts and clefts in Akan specifically, and then move on to the theoretical approaches that argue for parallelism between definite pseudoclefts and clefts, which I will refer to as the *semantic definite* accounts, following De Veaugh-Geiss et al. 2018.

4.2.4.1 The view from Akan

In a seminal paper on clefts in Akan, Boadi (1974) claims that definite descriptions with identity statements of the form '*Nipa no a* P *ne* α' ('The person who did *P* is α') are focus-marking constructions just like *nà*-clefts—and in fact, the one is derived from the other. The analysis Boadi (1974) proposed is as follows: Focus-marking in Akan requires undergoing a 'Focus Attachment Rule'. For *na*-clefts, the basic focus marker is attached to the SVO clause, e.g., [na Kofi baa ha], and the focus is copied from its base position to the left periphery, e.g., [Kofi na Kofi baa ha]; finally,

pronominalization in base position replaces the left-dislocated constituent, e.g., ɔ- in (30).

(30) a. [Kofi$_i$] na [ɔ$_i$-ba-a ha].
Kofi FM 3.SG-come-PST here
'*Kofi is the one who came here.*'
[ex. (1b) and translation from Ofori 2011]

For definite pseudoclefts the derivation is essentially the same: The underlying main clause is an identity statement with *y*ɛ 'is', e.g., [ɔ$_i$-baa ha *y*ɛ Kofi$_i$]. The basic focus marker *na* attaches to the underlying clause [na + ɔ$_i$-baa ha *y*ɛ Kofi$_i$]; and *ne* is formed by the fusion of the focus marker with the copula verb *y*ɛ. Finally, a headed relativizer, e.g., *onipa a* 'person who', is placed clause-initially. The result is as in (31a), and the variant in (31b) involves constituent swapping.

(31) a. [Onipa a ɔ$_i$-ba-a ha] ne [Kofi$_i$].
person who 3.SG-come-PST here FM Kofi
b. [Kofi$_i$] ne [onipa a ɔ$_i$-ba-a ha].
Kofi FM person who 3.SG-come-PST here
'*Kofi is the one who came here.*'
[ex. (1c) and ex. (12a) from Ofori 2011]

Ofori (2011), who similarly proposes a unified analysis of clefts and pseudoclefts, takes the opposite approach, instead analyzing *ne* as the basic form of the focus marker from which *na* is derived as a fusion of the *ne* focus marker and the relative pronoun *a* 'who'. Ofori (2011) makes the critical point that the morphophonological fusion of *na* and *y*ɛ proposed by Boadi (1974) is not language-internally motivated in Akan; by contrast, deriving *na* from the merging of *ne* and the relative marker *a* has parallels in the morphophonological reduction found in Akan possessive constructions. In Ofori's (2011) proposal, the basic focus sentence is a biclausal structure as in (32), in which the relative clause is headed by a generic noun phrase *onipa* coindexed with the proper name subject of the identity statement, *Kofi*, and the resumptive pronoun ɔ- in the relative clause.

(32) Basic focus construction (headed relative clause)
a. Kofi$_i$ ne onipa$_i$ a ɔ$_i$-ba-a ha.
Kofi FM person who 3.SG-come-PST here
'*Kofi is the person who came here.*'
[adapted from ex. (12a) in Ofori 2011]

For the *nà*-cleft, the construction is derived from the headless variant of (32), shown in (33). Crucially, regarding relativization in Akan Ofori (2011) observes

that relative clauses may separate the head noun from the relative clause when extraposed; furthermore, adjectives and verbs can also be relativized, whereby the relative pronoun is not headed by a noun phrase. Ofori (2011) presents (33) as an example of when *ne* selects a headless relative clause, which is derived via NP-deletion of the generic noun phrase *onipa* 'person' from the headed variant in (32). The morphonological fusion of the remaining contiguous elements *ne* + *a* results in the *nà*-cleft. (See also Titov 2019: §5, who discusses the relationship between clefts and restrictive relative clauses in Akan in detail, and similarly argues for *nà* being derived from the copula and relative pronoun.)

(33) Derived focus construction (headless relative clause)

a. Kofi$_i$ ne ~~onipa$_i$~~ a ɔ$_i$-ba-a ha.
Kofi FM who 3.SG-come-PST here

→ Morphophonological Fusion
Kofi$_i$ na ɔ$_i$-ba-a ha.
Kofi FM+who 3.SG-come-PST here
'*Kofi is who came here.*'

[adapted from ex. (12b) in Ofori 2011]

As a result, relative clauses which never have a generic noun head, such as adjectives and verbs, are predicted to only have the *na*-focus form, which is in line with the data (Ofori 2011: pp. 250–251). Furthermore, as Ofori (2011: p. 253) argues: "The potency of the present account over Boadi (1974) lies in the fact that we provide language-internal support for every position advanced in a way that renders the present account more credible. For the first time we are able to posit a single basic focus sentence for Akan and also are able to explain the *ne* and *na* focus sentence difference in a way that is consistent with the phonotactic and morphotactic principles of Akan."

4.2.4.2 Theoretical approaches to exhaustivity in definite pseudoclefts and clefts

Although differing in the details, the approaches of Boadi (1974) and Ofori (2011) share the view that definite pseudoclefts and *na*-clefts are underlyingly the same, and similar claims have been argued in the theoretical literature for other languages as well (among others, Akmajian 1970, Percus 1997, Büring & Križ 2013, Križ 2016, 2017; although see Destruel & De Veaugh-Geiss 2019 for diverging results for French between *c'est*-clefts and the definite pseudoclefts tested there). Such accounts are collectively referred to as the SEMANTIC DEFINITE accounts in De Veaugh-Geiss et al. 2018, a term which subsumes a diverse group of theoretical

analyses which conventionally-code exhaustivity as part of the meaning of definite pseudoclefts and clefts.

Under one approach, in which *it*-clefts are derived from an underlying definite description with identity statement (Percus 1997), the exhaustivity inference in clefts *is* the uniqueness or maximality presupposition of the definite description plus the identificational semantics. By contrast, in the most recent iteration of the homogeneity approach presented in Büring & Križ 2013 and Križ 2016, Križ (2017) proposes that "*it*-clefts semantically correspond to copular sentences with the cleft predicate turned into a (number-neutral) definite description" (Renans & De Veaugh-Geiss 2019: p. 4). Thus, a violation of exhaustivity is "conceptually the same as a so-called *homogeneity violation* in a sentence with a plural definite description" (Križ 2017: p. 2; see Renans & De Veaugh-Geiss 2019: §2.2 for a detailed summary of this approach). A crucial aspect shared by these SEMANTIC DEFINITE accounts is that exhaustivity/homogeneity in clefts and definite pseudoclefts is predicted to invariably arise regardless of context.

On the other hand, De Veaugh-Geiss et al. (2018), Destruel & De Veaugh-Geiss (2018, 2019), Zimmermann et al. (to appear) posit that clefts and definite pseudoclefts—at least anaphoric definite pseudoclefts—obtain an exhaustive interpretation discourse-pragmatically via the resolution of the anaphoric existence presupposition encoded in both sentence types (see also Pollard & Yasavul 2016 for clefts; cf. Horn 1981, 2014, who similarly derives exhaustivity from the cleft existential). That clefts have anaphoric potential has long been observed in the literature (see, e.g., Delin 1992). In example (34), for instance, De Veaugh-Geiss et al. (2018) argue that "on its only licit interpretation [. . .] the discourse referent introduced by the indefinite in the first sentence is anaphorically picked up and further specified in the cleft sentence"

(34) a. Judy was looking for somebody all afternoon.
b. It was her youngest daughter that Judy was looking for.
[ex. (21) in De Veaugh-Geiss et al. 2018]

Under this approach, depending on how the anaphoric existence presupposition is resolved, the discourse antecedent will be maximal with an exhaustive interpretation, or non-maximal/indefinite with a non-exhaustive interpretation. I will refer to this account as the DISCOURSE PRAGMATIC approach. Notably, exhaustivity under this approach is derived from the anaphoric potential of clefts and definite pseudoclefts, and not from the exclusion of the focal alternatives (cf. De Veaugh-Geiss et al. 2015, see also Titov 2019); nor a uniqueness/maximality presupposition (cf. Percus 1997); nor as part of a homogeneity-triggering construction (cf. Križ 2017).

	± *parallel def.pse.*	± *strength*
(A) *semantic definite*	+	+
(B) *discourse pragmatic*	+	−

Table 4.1 *Theoretical predictions regarding the expected strength of exhaustivity for the semantic definite and discourse pragmatic approaches to cleft exhaustivity.*

4.2.5 Theoretical predictions

Shown in Table 4.1 is a summary of the theoretical predictions for exhaustivity in clefts and definite pseudoclefts in light of the discussion above. The term *strength* is borrowed from De Veaugh-Geiss et al. (2018: p. 6), which they define in the following way: "*strength* is a cover term that refers to the overall robustness and systematicity of an exhaustivity inference. A robust inference is both obligatory and non-cancellable across all contexts for all speakers of a language group. The term systematicity is related to the notion of robustness but it refers specifically to the regularity of exhaustivity across experimental set-ups, experimental conditions, and also across speakers."[16]

As can be seen in the table, the SEMANTIC DEFINITE and DISCOURSE PRAGMATIC accounts are similar in predicting parallel behavior between clefts and definite pseudoclefts, but they differ in the strength of the exhaustive inference. Given that exhaustivity is conventionally-coded in the cleft structure in the semantic definite accounts, exhaustivity is predicted to be a strong inference which arises regardless of context. By contrast, the discourse pragmatic account predicts either a strongly exhaustive or non-exhaustive interpretation depending on how the existential presupposition is resolved; and thus, potential variation in participant responses is expected.

16 It should be noted by the reader that—for reasons which will soon become clear—the analysis presented in Section 4.3 is exploratory, and prior to running the experiments the theoretical predictions were as described in Section 2.1 in De Veaugh-Geiss et al. 2018. When the experiment on Akan was initially planned and implemented, the discourse pragmatic analysis had not yet been proposed. Moreover, for reasons of space, and since the analysis in this manuscript is post hoc, I have not included a discussion of the theoretical predictions of the pragmatic accounts of Horn (1981, 2014) and De Veaugh-Geiss et al. (2015) nor the semantic Inquiry-Terminating account of Velleman et al. (2012) as they might apply to Akan. Neither approach makes any particular prediction regarding parallels between clefts and pseudoclefts, which is of primary interest for the exploratory analysis here.

Of secondary interest is a hypothesis concerning subject- vs. object-focus *nà*-clefts in terms of the exhaustive inference.[17] As discussed above, the *in situ* strategy is claimed to be the prefered focus-marking strategy with object focus (Genzel 2013). Thus, one hypothesis is that object focus with the *nà*-cleft will result in a manner implicature. As Grice (1975: p. 46) wrote, the maxim of manner relates "not [. . .] to what is said but, rather, to HOW what is said to be said", and includes maxims such as "[a]void obscurity of expression" and "[b]e brief (avoid unnecessary prolixity)". In this view, object-clefts may elicit a strong exhaustivity effect when appearing *ex situ* since the speaker has 'gone out of her way' (to quote Horn 1981: p. 133) in using the cleft construction. By contrast, *nà*-clefts with subject focus will lack this potential implicature since the cleft construction is the default strategy.[18] Importantly, this implicature for non-subject focus is expected only to arise in *nà*-clefts, since no comparable subject/non-subject asymmetry appears to exist for the definite pseudocleft construction. Without further ado, I turn now to the experiments.

4.3 Experiments

The studies presented here for Akan are identical to the methods and design of the experiments described in De Veaugh-Geiss et al. 2018 for German. In a nutshell, in the studies on German, De Veaugh-Geiss et al. did not find evidence of a difference between *es*-clefts and definite pseudoclefts of the form *Derjenige, der* . . . 'The one who . . . '. Moreover, for the cleft and pseudocleft conditions German participants clustered into two 'exhaustivity' groups: those who treated both sentence types as strongly exhaustive; and those who treated both as non-exhaustive. The authors interpret the results as being compatible with exhaustivity not being encoded in clefts and definite pseudoclefts; rather, De Veaugh-Geiss et al. 2018 propose the inference is discourse-pragmatically derived via the resolution of the anaphoric existence presupposition (Pollard & Yasavul 2016). For details on the German version of the experiment, I refer the reader to De Veaugh-Geiss et al. 2018. Here I will focus on

17 I refer to this manipulation as secondary since—given various practical limitations while preparing and running the experiments in Ghana—it was ultimately decided that the subject/non-subject alternation be a between-item manipulation in order to reduce the number of items and lists needed overall, effectively reducing statistical power; nevertheless, subject/object focus was systematically distributed across the experimental trials in order to compare the two grammatical arguments. See Section 4.3.1 'Methods & Design' for more details about the materials, and Appendix D.2 for a presentation of all items.

18 Cf. French *c'est*-clefts, which are similarly the default strategy for subject-focus, discussed at length in Destruel & De Veaugh-Geiss 2018. If the proposal in Destruel & De Veaugh-Geiss 2018 is on the right track, however, since *nà*-clefts cannot be used to signal broad focus, there is no expectation that subject *nà*-clefts will exhibit a weak exhaustivity inference in Akan, in contrast to French.

the details of the Akan experiment, before directly comparing the results of this study with the data for German from De Veaugh-Geiss et al. 2018.[19]

4.3.1 Methods & Design

4.3.1.1 Participants

For the Akan version of the experiments: In Experiment I (verifier) 29 participants (male: 9, female: 20; mean age: 22.48, age range: 18–32) were tested, and in Experiment II (falsifier) 29 participants (male: 9, female: 20; mean age: 21.34, age range: 18–25) were tested; note that the participants in Experiment I were distinct from the participants in Experiment II.[20] Participants, mostly undergraduate and graduate students, were recruited at the University of Ghana at Legon in February 2017 over a period of two weeks. Participants were compensated 10 Ghanian cedis for their participation. For the German version of the experiments there were 32 participants in Experiment I (verifier)[21] and 32 participants in Experiment II (falsifier); again, the participants in each experiment were distinct from one another (see De Veaugh-Geiss et al. 2018 for further details).

4.3.1.2 Materials

All materials, including the instructions and target/filler items, were constructed by the author together with several Akan native-speaker students at the Linguistics Department at the University of Ghana at Legon. See Appendix D.2 for all target stimuli.

There were four sentence types with exhaustivity inferences tested, two control conditions and two target conditions. The control conditions included (i) sentences with the exclusive particle *nkóáá* 'only', illustrated in (35a), and (ii) a non-cleft SVO sentence with *in situ* focus, illustrated in (35b); the target conditions included (i) a definite pseudocleft construction (definite description with an identity statement),

19 German data and target stimuli available at https://static.semprag.org/sp.11.3s.zip; Akan data available at https://gitup.uni-potsdam.de/deveaugh/public-repository-for-the-paper-na-cleft-non-exhaustivity-variability-in-akan.

20 A total of 4 participants were removed from the data set for Akan given unexpected response behavior in the exclusive control condition: in the manipulation which should lead to the sentence with the exclusive being judged 'not true', namely when exhaustivity was falsified, 2 participants in Experiment I (verifier) and 2 participants in Experiment II (falsifier) judged the exclusive sentence as 'true' a majority of the time (3/4 'true' judgments). See Appendix D.1 for details.

21 This is after data preparation. As De Veaugh-Geiss et al. (2018: Fn. 8) write: "There were 33 participants in Experiment I [in German], but one participant was removed for erratic judgments on the exclusive control condition: 2 'continues,' 2 'false,' and 3 'true' judgments as well as one missing data point at Box 2 due to already having made a judgment at Box 1".

illustrated in (35c), and (ii) a *nà*-cleft, illustrated in (35d). Note that for the auditory stimuli, in order to ensure that focus was on the intended constituent, e.g., on *Anan* in (35), the native-speaker who recorded the stimuli read all sentences as an answer to a narrow-focus *wh*-question, e.g., *Who did Yaa kick out?* All examples below are with object focus, with the focus indicated by bracketing below for presentational purposes. The manipulation of the grammatical argument involved having the entity in brackets—which was invariably one of four fictional roommates, discussed below—be either the grammatical subject or object. (The materials for the German version of the experiment were directly comparable in form though differing in the lexicalizations and, importantly, the grammatical argument manipulation; again, see De Veaugh-Geiss et al. 2018 for details.)

(35) **Example** – Item 124

a. [Anan]$_F$ nkoaa na Yaa pamoo no.
Anan only FM Yaa kicked.out 3.SG.OBJ
'*It is only* ANAN *who Yaa kicked out.*' (*Exclusive*)

b. Yaa pamoo [Anan]$_F$.
Yaa kicked.out Anan
'*Yaa kicked out* ANAN.' (*SVO*)

c. Nipa no a Yaa pamoo no ne [Anan]$_F$.
person the who Yaa kicked.out 3.SG.OBJ is Anan
'*The person who Yaa kicked out is* ANAN.' (*Def.Pse.*)

d. [Anan]$_F$ na Yaa pamoo no.
Anan FM Yaa kicked.out 3.SG.OBJ
'*It is* ANAN *who Yaa kicked out.*' (*Cleft*)

Since there were two sentence forms given the subject/object manipulation, the description here will be broken down into two parts. I will start with objects. The schema for the auditory stimulus in each sentence type condition with object focus was as follows; note that SBJ stands for subject, OBJ stands for object, V stands for verb, RPRON stands for resumptive pronoun, and PRT stands for particle.[22]

(36) EXCLUSIVE: [ROOMMATE]$_F$ only *nà* SBJ V RPRON (PRT)
SVO: SBJ V [ROOMMATE]$_F$ (PRT)
DEF. PSE.: person the who SBJ V RPRON (PRT) is [ROOMMATE]$_F$

22 With the term particle I wish to remain neutral as to the grammatical category of this clause-final element, which sometimes appears to be similar to an adverb and sometimes an inherent verbal complement. Note that half of the items appeared without a particle (as in example (35) above) and half with a particle (for an example, see the word *mu* 'inside' for Item 101 in Appendix D.2). Although in some cases the lexicalization of the verb is the same across multiple target items, the additional verbal particle changed the verb's meaning (cf. the English verb *take* in *take something/someone out* vs. *take something/someone on*).

CLEFT: [ROOMMATE]$_F$ *nà* SBJ V RPRON (PRT)

The sentences with object focus had the following characteristics:

- Object: one of four fictional roommates (*Anan*, *Kofi*, *Kwaku*, *Kwame*).
- Subject: one of 16 unique proper names (*Yaa*, *Adwoa*, *Kumi*, etc.).
- Verb tense: 8 target items in past tense and 8 in present tense.

Importantly, by having four fictional 'roommates', the domain of focus alternatives was held constant to avoid issues related to domain widening. Moreover, since the object was animate, a resumptive pronoun appeared either in the main clause when the object was fronted with *nà* or in the relative clause of the definite pseudocleft condition. Finally, in order to avoid any 'definiteness' effects, neither the cleft nor the definite pseudocleft sentences had the optional determiner-like element *nó* in the right periphery of the clause (see Section 4.2.2).

The schema for the auditory stimulus in each sentence type condition with subject focus was as follows:

(37)
EXCLUSIVE: [ROOMMATE]$_F$ only *nà* RPRON-V OBJ (PRT)
SVO: [ROOMMATE]$_F$ V OBJ (PRT)
DEF. PSE.: person the who RPRON-V OBJ (PRT) is [ROOMMATE]$_F$
CLEFT: [ROOMMATE]$_F$ *nà* RPRON-V OBJ (PRT)

Note the following characteristics:

- Subject: one of four fictional roommates (*Anan*, *Kofi*, *Kwaku*, *Kwame*).
- Object: incorporated bare NP.
- Verb tense: 8 target items in past tense and 8 in present tense.

All 32 items in the target/control trials and 32 items in the filler trials had unique semantic meanings, and all proper names (except the four fictional roommates) and all objects had unique lexicalizations per trial. There were 64 trials in total.

4.3.1.3 Procedure

The experiments used a sentence-picture verification task in an incremental information-retrieval paradigm (Conroy 2008, Franke, Schlotterbeck & Augurzky 2016). The procedure was identical to that described in De Veaugh-Geiss et al. 2018, 2017 (German; see §3) as well as in Destruel & De Veaugh-Geiss 2019 (French; see §3) and Zimmermann et al. to appear (English and Hungarian; see §3); however, the post hoc exploratory analysis presented here differs from those studies.

In the instructions, participants were introduced to four fictional roommates, and participants were told that these four roommates will undertake various activities

Auditory Stimulus: *Anan na Yaa pamoo no.* 'It is Anan who Yaa kicked out.'

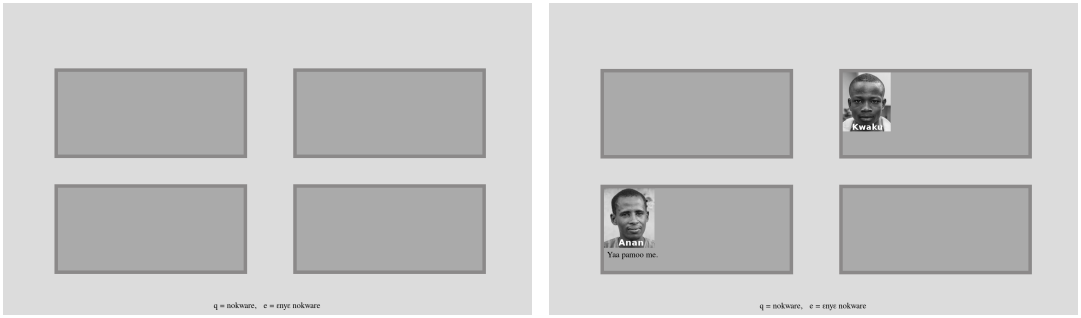


Figure 4.1 *(i) Start of each trial.* *(ii) Uncovering the 2nd box.*

together. At the start of each trial, a computer screen showing four covered boxes was presented while the stimulus played in the participants' headphones. The screen looked as in the left image in Figure 4.1. After hearing the auditory stimulus, participants were asked to uncover as many boxes as necessary to decide if the sentence they heard was *nokware* 'true' or *ɛnyɛ nokware* 'not true'. Each box contained a photo of one of the roommates[23] and a written first person statement about which activity this roommate did, as in the right image in Figure 4.1. The experimental software was programmed in Python (PSFL: GNU/Linux v.3.4.2) using the PyGame module (LGPL: v.1.9.2a0, Shinners 2011) by Edgar Onea, which allowed full control over the experimental setup.

Both experiments began with three practice trials using stimuli that were unrelated to the experimental materials in order to ensure that the participants understood that they are supposed to uncover only as many boxes as necessary. When participants uncovered too many boxes in the practice trials, they were reminded by the researcher observing them not to do so.

23 The photos for the fictional roommates are used according to the terms of the Creative Commons BY-SA license: https://creativecommons.org/licenses/by-sa/2.0/. The only modifications made to them were (i) cropping the image to fit the screen and (ii) converting the color photos to black-and-white. The four photos were taken by Mark Fischer (from the 'Abuja Street Portrait' and 'N'Djamena Street Portrait' series) and are available for download at the following URLs:

- https://www.flickr.com/photos/fischerfotos/16360022790/
- https://www.flickr.com/photos/fischerfotos/22986270194/
- https://www.flickr.com/photos/fischerfotos/23519903990/
- https://www.flickr.com/photos/fischerfotos/23484547082/

For 23 participants (Experiment I) and 22 participants (Experiment II) out of the total 29 participants per experiment, the boxes were uncovered by moving the cursor over them with the mouse, and after entering a box the cursor could not exit for at least 2000 ms. This was done to disincentivize unnecessary uncovering of the remaining boxes. When the cursor was eventually moved out of the box, the text inside disappeared. Participants were free to choose which box they uncovered next, but they did not know that their choice had no influence on the order of information they received—the order was pre-determined by the experiment software. Participants were instructed to press one of two designated keys on the keyboard for either 'true or 'not true' once they had enough information to make a decision. Once participants made a judgment the boxes onscreen were recovered and the next target or filler item played in their headphones.

For the final 6 participants in Experiment I and 7 participants in Experiment II the procedure was slightly different. Anticipating the discussion to come, almost all of the participants had behaved unexpectedly in the SVO and EXCLUSIVE control conditions: namely, a response strategy was employed of first uncovering all boxes before making a truth-value judgment regardless of the information presented on screen. Thus, to strongly disincentivize an uncover-first-judge-later strategy, for these 13 participants the controls were changed to be keyboard-driven (i.e., the space bar was used to uncover boxes) and, crucially, the delay between uncovering was increased to 5000 ms. This increased delay doubled the total running time of the experiment. Although the extended delay influenced to a slight degree when participants made a judgment, many participants still uncovered all boxes before making a decision; furthermore, and most importantly in light of the final judgments made, the reasons for further uncovering remain unclear (cf. Bombi & De Veaugh-Geiss 2018 for a discussion of issues which arise when naively employing quantitative methods in different cultural settings).

Experiment I (verifier) In Experiment I, the contextual information provided in Box 2 (early manipulation) *verified* the canonical meaning or prejacent of the target stimulus (i.e., the sentence equivalent to the SVO form); hence, I will refer to this experiment as 'verifier' in order to distinguish it from Experiment II, which used the exact same design but had a different manipulation at Box 2. Consider, for instance, the exclusive control stimulus with object focus in (35a), repeated in (38).

(38) *Auditory Stimulus*

Anan nkoaa na Yaa pamoo no.
Anan only FM Yaa sack.PST 3.SG.OBJ
'*It is only* ANAN *who Yaa kicked out.*' [Target Item 124]

In this case, Box 2 showed one of the fictional roommates, here Anan, stating *Yaa kicked me out*, shown in Figure 4.1, thus verifying the canonical meaning of the sentence *Yaa kicked out Anan*.

For the later boxes, there was an additional late manipulation, ±EXH. Namely, in half of the trials one of the remaining two roommates falsified exhaustivity (–EXH), which was equally distributed across Box 3 and Box 4. The other half of the trials were compatible with an exhaustive interpretation of the auditory stimuli (+EXH), not shown here: in this case, of the four roommates only Anan in Box 2 states that Yaa kicked him out; that is, the sentence will be true for all sentence types. An example of a full trial with a verifier in Box 2 (+VER) and a falsifier in Box 4 (–EXH) for auditory stimulus (38) is shown in (39) below.

(39) Full trial for Experiment I (verifier)
Auditory Stimulus: '*It is only* ANAN *who Yaa kicked out.*'

Box 1: Kwaku

Yaa titii me ho.
Yaa scratch.PST 1.SG.OBJ around
'*Yaa scratched me.*'

Box 2: Anan (+VER)

Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Box 3: Kofi

Yaa hyiraa me.
Yaa bless.PST 1.SG.OBJ
'*Yaa blessed me.*'

Box 4: Kwame (–EXH)

Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Experiment II (falsifier) In Experiment II, a mirror of Experiment I, the contextual information provided in Box 2 (early manipulation) *falsified* the exhaustivity inference of the target stimulus; hence, I will refer to this experiment as 'falsifier' in order to distinguish it from Experiment I (verifier), which used the exact same design but had a different manipulation at Box 2. For instance, consider any of the targets sentences in (35), but now in Box 2 Kwame (instead of Anan, who has not been revealed yet) is stating *Yaa kicked me out*, violating the exhaustivity inference that Yaa kicked out nobody other than Anan. This is shown in example trial (40) below.

For Box 3 or Box 4, there was an additional manipulation, ±CAN. Namely, the canonical meaning or the prejacent of the target stimulus was verified in half of the cases (+CAN); for instance, Anan says Yaa kicked him out, shown in Box 4 in (40) below. In the other half of the trials, the canonical meaning was not verified (–CAN), not shown here; in this case, the sentence is not true for all sentence types, since Anan says that Yaa did something other than kick him out. An example of a full trial

with a falsifier in Box 2 (+FAL) and a verifier in Box 4 (+CAN) for auditory stimulus (38) is illustrated in (40).

(40) Full trial for Experiment II (falsifier)
Auditory Stimulus: '*It is only* ANAN *who Yaa kicked out.*'

Box 1: Kofi Yaa hyiraa me. Yaa bless.PST 1.SG.OBJ '*Yaa blessed me.*'	**Box 2: Kwame (+FAL)** Yaa pamoo me. Yaa sack.PST 1.SG.OBJ '*Yaa kicked me out.*'
Box 3: Kwaku Yaa titii me ho. Yaa scratch.PST 1.SG.OBJ around '*Yaa scratched me.*'	**Box 4: Anan (+CAN)** Yaa pamoo me. Yaa sack.PST 1.SG.OBJ '*Yaa kicked me out.*'

In the German version of the experiment, the distribution of items for Box 2 as well as for Box 3/4 was fully balanced: for instance, Item 2 appeared both in the [+VER, +EXH] and [+VER, –EXH] conditions in Experiment I (verifier), as well as in the [+FAL, +CAN] and [+FAL, –CAN] conditions in Experiment II (falsifier). This was not so in the Akan versions of the experiments. For Akan, the distribution of items was such that the late manipulations were fixed per item: for example, the late manipulation for Item 102 always appeared in the [+EXH] condition in Experiment I (verifier) and in the [–CAN] condition in Experiment II (falsifier). This was done for two reasons: First, the outcome variable was intended to be whether participants made a truth-value judgment at the early manipulation at Box 2—which was indeed the same for all experiments and languages—and the late manipulations at Box 3/4 were controls that participants understood the logic of the experiment (see De Veaugh-Geiss et al. 2018: §3.4 for discussion). Thus, for the intended analysis, fixing the late manipulations was not critical. Second, there were various on site limitations while conducting the experiments (cf. Bombi & De Veaugh-Geiss 2018); thus, some changes were made to simplify the experiment, and having the late manipulations fixed per item was one way to reduce the number of lists needed for the Latin square distribution. As a result, for the exploratory analysis here not all combinations are represented for all items in the data in Akan.[24]

24 For the sake of being explicit: this design change unfortunately means that some items will have to be left out of the final analysis. For instance, in Akan Item 102 did not appear in the combinations when the canonical meaning is verified but exhaustivity violated—that is, [+VER, –EXH] in Experiment I (verifier) or [+FAL, +CAN] in Experiment II (falsifier)—although it is precisely these combinations that will be of primary interest here. Furthermore, although many items are a within-experiment

Thus, in the original design of the experiment, the dependent variable was whether a judgment was made at Box 2 or whether participants continued uncovering Box 3 and Box 4. Although Experiment I (verifier) and Experiment II (falsifier) are mirrors of each other, manipulating the information at Box 2 served a crucial purpose. As discussed in De Veaugh-Geiss et al. 2018: §3.4, running both experiments side-by-side potentially allowed the authors to detect a hypothetical class of not-at-issue inferences they refer to as *not-immediate* inferences. They posit that such inferences might not come to mind or be ignored by a speaker when the inference is not immediately necessary for evaluation of truth or falsity. This could be the case for clefts, for instance, when Box 2 verifies the at-issue assertion corresponding to the canonical meaning, and the exhaustivity inference has yet to be confirmed/violated. However, when confronted with an explicit violation of a not-immediate inference, such as when Box 2 falsifies the not-at-issue exhaustive meaning of the cleft, the inference can no longer "be disregarded" (De Veaugh-Geiss et al. 2018: p. 30). Nevertheless, clefts and pseudoclefts did not exhibit response patterns suggesting the inference was a not-immediate one.

As discussed in the following section, what is analyzed here is not the proportion of judgment vs. continue responses at Box 2 as per the intended design, but the final truth-value judgments regardless of when that judgment was made. The reason for that will become clear shortly.

4.3.1.4 Motivation for the exploratory analysis

As previously mentioned, the participants in the Akan version of the experiment generally exhibited an uncover-first-judge-later response strategy, a characterization which in this section I will explore in detail. First, I will take a look at the results for the control conditions at Box 2, in which either the canonical meaning was verified (Experiment I) or the exhaustivity inference falsified (Experiment II). In these early manipulations, the German and Akan data appear to differ in crucial ways. After this I will present a different way of looking at the data—taking into account both the early and late manipulations together—which suggests participants were engaged in the task. Following this, I will present the linking hypotheses and research questions for the post hoc exploratory analysis presented in Section 4.3.2 'Results'.

Control conditions at Box 2 At Box 2, results differ quite drastically between Akan and German in the SVO and exclusive control conditions, shown in Figure 4.2. In this graph, the proportion of 'continues' (i.e., when participants continued

manipulation, relevant for the random effects structure of the statistical model, not all items are; see Appendix D.1 for an overview of the structure of the data used in the analysis.

Figure 4.2 *A comparison of the proportion of 'continues' at Box 2 in the* SVO *(Experiment I) and* EXCLUSIVE *(Experiment II) control conditions in Akan and German; note that the responses here were made before uncovering Box 3/4. Since enough information has been revealed for the control conditions at Box 2 to make a truth-value judgment, it was expected that participants would not continue uncovering Box 3/4, contrary to what is found in Akan.*

to uncover Box 3/4) is plotted.[25] The predictions for the control conditions were as follows: For the SVO condition in Experiment I (verifier), since exhaustivity is a non-truth-conditional inference it was expected that participants would make a 'true' judgment a majority of the time—i.e., the canonical meaning has been verified at this point—and *not* continue uncovering Box 3/4. By contrast, for the exclusive condition in Experiment II (falsifier), since exhaustivity is part of the truth-conditions it was expected that participants would make a 'not true' judgment—after all, the exhaustivity inference has been falsified—and *not* continue uncovering a majority of the time.

25 In the Akan version of the experiment, recall that the delay between uncoverings was 2000 ms for 45 participants and 5000 ms for 13 participants; however, the proportions shown here are averaged over both groups. To note, there was in fact a small difference in the proportion of continues made at Box 2 for the participants with the increased delay. Specifically, in Experiment I (verifier) the focus condition elicited 66% continues (29/44) and in Experiment II (falsifier) the exclusive condition elicited 47% continues (24/51)—i.e., averages which are just slightly lower than that found in Figure 4.2. Nevertheless, this difference is not relevant in light of participants' final truth-value judgments.

As can be seen, the German participants choose to make a judgment a majority of the time, seen in the low proportion of 'continues' in Figure 4.2, whereas the Akan participants largely choose to continue uncovering in these conditions. A naive interpretation of these results is as follows: SVO sentences with *in situ* focus in Akan have a stronger exhaustivity inference than in German, since the Akan participants in Experiment I (verifier) choose to keep uncovering Box 3/4 in order to check that the exhaustivity inference holds. Moreover, the exclusive *ńkóáá* 'only' is not nearly as exhaustive as the German counterpart *nur* 'only', since despite the exhaustivity violation the participants in Experiment II (falsifier) continued uncovering in order to check whether the canonical meaning is verified.

This interpretation, however, would fly in the face of the literature, in which it is claimed that SVO sentences with *in situ* focus in Akan are only weakly exhaustive at best, and, moreover, the exclusive particle in Akan is typically analyzed as on a par with exclusive particles in other languages (see Section 4.2.3 'Exhaustivity inferences across sentence types' for more information). Moreover, by looking at the data from a different perspective it is clear that the above interpretation is incorrect. That is, instead of analyzing what happens at Box 2, one can look at the final truth-value judgments irrespective of when that judgment was made—taking into consideration both the early and late manipulations—and a wholly different cross-linguistic pattern is found.

A different perspective: final truth-value judgments Looking at *final* truth-value judgments made while taking into consideration both the early and late manipulations—and not concentrating on the judgment vs. continue responses at Box 2—parallels between the Akan and German versions of the experiments emerge. Note that there are four possible combinations for the Box 2 plus Box 3/4 manipulations (Experiment I: +VER, ± EXH; Experiment II: +FAL, ± CAN). Here I will present the two manipulations which are predicted to elicit invariably 'true' judgments and invariably 'not true' judgments for all sentence types. By considering the response patterns in these two manipulations first, I can be confident that participants were in fact sensitive to the information presented in the experimental trials.

There are two manipulations which validate that participants were engaged in the task and paying attention to the truth or falsity of the sentences in context. In Experiment I, when participants find that the canonical meaning is verified at Box 2, and—in the event they continue to uncover boxes—they discover that exhaustivity holds across Box 3 and Box 4 (i.e., [+VER, +EXH]), it is predicted that the sentence will be judged 'true', regardless of sentence type. In fact, this is exactly what was found for both languages (modulo noise). By contrast, in Experiment II, when participants find that the exhaustivity inference is falsified at Box 2, and—in the event they continue to uncover boxes—they discover that the canonical meaning

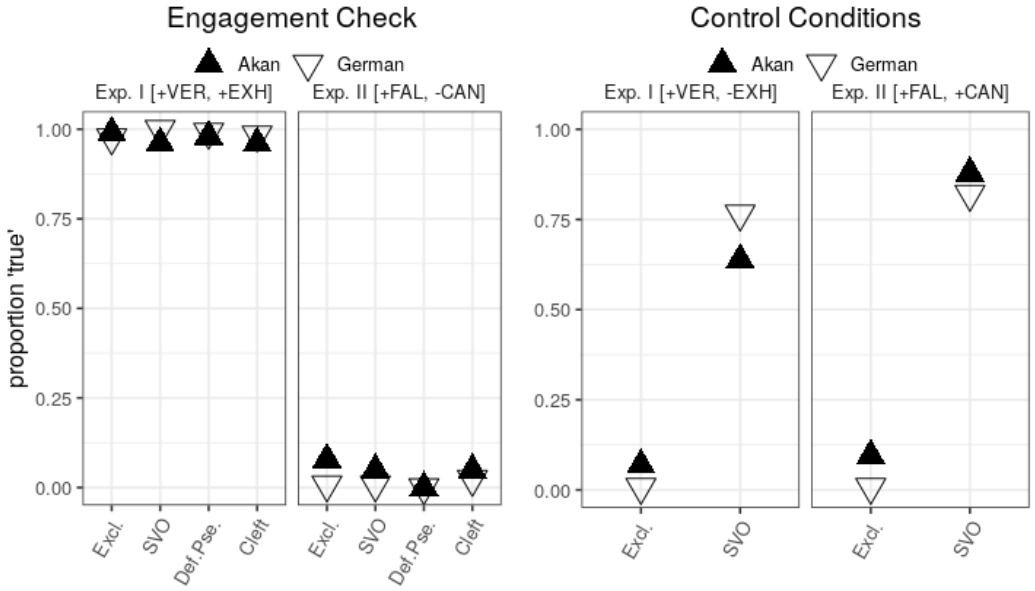


Figure 4.3 *Left: A comparison of Akan and German in the final truth-value judgments (0 = not true, 1 = true) when the early and late manipulations are predicted to elicit either true* ([+VER, +EXH]) *or not true* ([+FAL, –CAN]) *judgments for all sentence types, irrespective of the box at which the judgment was made. Right: Results for the control conditions in the critical combinations* [+VER, –EXH] *and* [+FAL, +CAN]*, i.e., when the canonical meaning holds but exhaustivity is violated.*

does not hold in either Box 3 or Box 4 (i.e., [+FAL, –CAN]), it is predicted that the sentence will be judged 'not true', regardless of sentence type. Again, this pattern is precisely what was found for both languages (modulo noise). These results are found in the left graphs in Figure 4.3.

A note on the results: The final truth-value judgment had only two possible outcomes, 'true' or 'not true'; in the plots in Figure 4.3, the y-axis shows the proportion of 'true' judgments. One may notice that in some cases the Akan results are slightly below ceiling or above floor, in particular in comparison to the German data. I posit that this is due to the differing response strategies of the participants: whereas the German participants made their judgment immediately at the box in which the critical information was provided (see Section 3.4 in De Veaugh-Geiss et al. 2018), the Akan participants uncovered all boxes first and only later made a judgment. This uncover-first-judge-later strategy arguably increases the cognitive load of the task and, in turn, it is not surprising that there is more noise in the data.

Linking hypotheses and research questions For the final analysis in Section 4.3.2, I will focus on the two critical combinations which are identical in informational import when looking at the unordered combinations of early and late information, namely [+VER, –EXH] and [+FAL, +CAN]—i.e., when the canonical meaning holds but exhaustivity is violated. When analyzing the final truth-value judgments, there are only two response choices, 'true' and 'not true'. The linking hypothesis is, if exhaustivity can be violated and still be judged 'true', then it must not be coded as part of the meaning of the sentence. Thus:

- If exhaustivity is conventionally-coded as part of the meaning of the sentence (semantic) ⇒ the target sentence will generally be judged 'not true' in the critical combinations since the exhaustivity inference has been violated. The control condition is the exclusive condition. Moreover, I am assuming that violations of presuppositions will similarly lead to 'not true' judgments; see discussion in De Veaugh-Geiss et al. 2018.

- If exhaustivity is a non-conventionally-coded inference (pragmatic) ⇒ the target sentence will generally—but not necessarily—be judged 'true' in the critical combinations since the exhaustivity inference can be cancelled; that is, participants will make their judgment primarily depending on the verification of the canonical meaning of the sentence (although strengthening effects may still apply). The control condition is the SVO condition with *in situ* focus.

By analyzing the final truth-value judgments in the critical combinations for clefts and definite pseudoclefts, the research questions I address in this paper are presented in Q1–Q3 below.

Similar to the semantic-pragmatic division found in the theoretical literature on *it*-English clefts (see De Veaugh-Geiss et al. 2018: §2.1 for an overview), it has been claimed both (i) that exhaustivity in *nà*-clefts is conventionally-coded and thus obligatory in unembedded contexts (e.g., presuppositional in Grubic, Renans & Duah 2019); and (ii) that it is not conventionally-coded in the cleft structure and thus not obligatory (Titov 2019). Recent literature on *es*-clefts and definite pseudoclefts in German (among other languages) argues that exhaustivity is a derived inference in both sentence types (e.g., De Veaugh-Geiss et al. 2018). Indeed, it is theoretically possible that German and Akan encode exhaustivity in clefts/pseudoclefts differently: e.g., exhaustivity in German clefts is a discourse pragmatic phenomenon, as argued in the DISCOURSE PRAGMATIC account in De Veaugh-Geiss et al. 2018, while Akan clefts encode exhaustivity as part of their conventional meaning, e.g., as modelled in the DISCOURSE SEMANTIC approaches. In fact, comparable cross-linguistic variation for cleft exhaustivity has been argued by Grubic, Renans & Duah (2019) for

Ngamo (West Chadic) vs. Akan/Ga (Kwa), with the latter argued to be a semantic inference, and the former a pragmatic inference; cf. clefts in St'át'imcets and Nłeʔkepmxcin, which are also claimed to lack a 'semantic' exhaustivity inference (Davis, Matthewson & Shank 2004, Koch & Zimmermann 2010). Similar semantic variation is found, for instance, in the existence presupposition of clefts (see, e.g., Davis, Matthewson & Shank 2004, Koch & Zimmermann 2010, Grubic, Renans & Duah 2019 for the lack of an existential for clefts in St'át'imcets, Nłeʔkepmxcin, and Ngamo; cf. clefts in German, Akan, English, etc., which are all claimed to encode existence as a presupposition) as well as other semantic phenomena such as third-person pronouns and determiners (Matthewson 2008). Furthermore, assuming the cleft and pseudocleft constructions in Akan are syntactically and semantically related, the two should exhibit similar exhaustivity effects. Thus, the first question is as follows:

Q1 *Cross-linguistic variability*: Are *nà*-clefts and definite pseudoclefts of the form *Nipa no a* ... 'The person who ...' strongly exhaustive in Akan, in particular in comparison to their counterparts in German? If the answer is yes, then this should be found in a main effect of language, and a positive result may provide evidence against taking a unified cross-linguistic analysis.

It is possible, though, that the results for clefts and pseudoclefts in Akan are compatible with a pragmatic analysis of exhaustivity similar to that proposed for German. Nevertheless, object *nà*-clefts may come with a manner implicature which strengthens the exhaustive inference—i.e., resulting in more participants perceiving clefting for object focus as strongly exhaustive—compared to subject clefts; by contrast, there is no reason to expect that definite pseudoclefts in Akan will give rise to a similar implicature for object focus. Thus:

Q2 *Inter-argument/inter-sentential variability*: Are Akan object clefts more strongly exhaustive than (i) subject clefts as well as (ii) definite pseudoclefts (for both grammatical arguments)? If the answer is yes, this should be found in an interaction of sentence type and argument.

Finally, given the by-participant results reported in De Veaugh-Geiss et al. 2018, another question that is of interest here is as follows:

Q3 *Inter-speaker variability*: If exhaustivity in Akan clefts/pseudoclefts is not a conventionally-coded inference, do participants cluster into 'exhaustivity' groups as was found in German? If the answer is yes, then this should be seen in a bimodal distribution when looking at individual response patterns.

I turn now to the results.

		Akan	*German*
EXCLUSIVE	*Exp. I*	8/116 'true' (7%)	1/128 'true' (<1%)
	Exp. II	11/116 'true' (9%)	1/128 'true' (<1%)
SVO	*Exp. I*	74/116 'true' (64%)	98/128 'true' (77%)
	Exp. II	102/116 'true' (88%)	105/128 'true' (82%)
DEF. PSE.	*Exp. I*	41/116 'true' (35%)	61/128 'true' (48%)
	Exp. II	52/116 'true' (45%)	54/128 'true' (42%)
CLEFT	*Exp. I*	49/116 'true' (42%)	60/128 'true' (47%)
	Exp. II	45/116 'true' (39%)	59/128 'true' (46%)

Table 4.2 *A descriptive summary of the number of 'true' judgments out of the total number of judgments per condition for Akan and German in the critical combinations* [+VER, –EXH] *and* [+FAL, +CAN]*, i.e., when the canonical meaning holds but exhaustivity is violated.*

4.3.2 Results

Here I will report the final truth-value judgments for the critical combinations, namely, when the canonical meaning holds but exhaustivity is violated (i.e., Experiment I [+VER, –EXH] and Experiment II [+FAL, +CAN]). When looking at the results for these combinations of early and late manipulations, differences between the sentence types emerge: whereas the SVO and EXCLUSIVE control conditions elicited a majority of 'true' and 'not true' judgments, respectively, CLEFTS and DEFINITE PSEUDOCLEFTS show a different—but parallel—pattern, and this across both languages.

4.3.2.1 Critical combinations

I refer to the combinations of early and late manipulations here as 'critical' since it is in these combinations that a difference between sentence types in terms of exhaustivity may emerge. Specifically, when the canonical meaning is verified but the exhaustivity inference is falsified, the final judgment of truth or falsity of the sentence will depend on the semantic/pragmatic encoding of exhaustivity in the constructions under consideration.

I will start with the control conditions, which is plotted in the right graphs in Figure 4.3 (on page 159 above). In Akan, the EXCLUSIVE condition elicited few 'true' judgments when exhaustivity was falsified (< 9%), and a similar pattern was found in German, although participants there made close to zero 'true' judgments

(<1%). A brief note on these results: It is, perhaps, not unexpected that the data for exclusives in German show close to floor results, whereas in Akan there is more noise. Recall that Akan participants employed an uncover-first-judge-later response strategy. As mentioned previously, this strategy will increase the demands of the task as more information must be stored in memory over a longer period of time; by contrast, German participants immediately stopped uncovering at the box with the falsifier. Nevertheless, the general response pattern for exclusives is as expected in both languages. Moreover, for the SVO sentences, the opposite pattern was found, also as expected. That is, in Akan participants made a 'true' judgment a majority of the time (64%–88%) in the SVO condition when exhaustivity did not hold, although to some extent less so in Experiment I than in Experiment II, and similar results were found in German (77%–82%).

As for the primary conditions of interest, in Akan the DEFINITE PSEUDOCLEFT elicited, on average, low but nevertheless mid-range 'true' responses (35%–45%), and in German a similar pattern was found (42%–48%). Moreover, results for clefts were comparable across languages: Akan *nà*-CLEFTS and German *es*-CLEFTS elicited 'true' responses almost half of the time (39%–47%), similar to their definite pseudocleft counterparts. See Table 4.2 for an overview of the descriptive results for all sentence types.

Bayesian analysis For an exploratory analysis of the data for clefts and definite pseudoclefts,[26] a binomial mixed logistic regression analysis in a Bayesian framework was conducted; see, e.g., Kruschke & Liddell 2018 and Vasishth et al. 2018 for a general introduction to Bayesian modelling. The statistics software *R* (v. 3.6.1, GPL-2 | GPL-3; R Core Team 2019) with the *brms* package (v. 2.9.0, GPL >= 3; Bürkner 2017, 2018) was used, which provides an interface to fit Bayesian models using Stan (New BSD License; Stan Development Team 2018). 122 participants

26 In order to focus on the primary conditions of interest, I report a model fit only on a subset of the data, that is, the responses for clefts and definite pseudoclefts. In an analysis not reported here, with a model fit to represent the full factorial design of the experiment (i.e., including all four sentence types), no differences were found for the cleft and definite pseudocleft conditions, and thus no conclusions differ from those discussed in this manuscript. There were nonetheless some results of interest. Using Helmert contrasts for the four sentence type comparisons, robust effects were found (i) when comparing the exclusive condition to the average of the SVO, cleft, and definite pseudocleft conditions; and (ii) when comparing the SVO condition to the average of clefts and definite pseudoclefts. Both of these results are in line with claims made in the literature for clefts in other languages (see Section 4.1). That said, there were some other robust effects beyond the scope of this paper. For instance, a difference between exclusives across the two languages was found, contrary to theoretical predictions. The near floor results for exclusives in German (<1% 'true' judgments), however, would likely make any statistical comparison between the two languages robust. Nevertheless, I will leave further analysis of this data for future research.

were included in the analysis (Akan: 29 participants per experiment; German: 32 participants per experiment; in total: 122 participants). Thus, when looking at the critical combinations of early and late manipulations, there are a total of 4 judgments per participant for each of the 2 sentence types, for 976 data points (122 * 4 * 2) in total.

Sum contrasts were used for all comparisons (Schad et al. 2019): the within-participant, within-item factor sentence type *ST* (DEFPSE 1/2, CL –1/2); the between-participant, between-item factor language *LANG* (AK –1/2, GE 1/2); as well as the between-participant and (partially) within-item factor experiment *EXP* (EXP.I 1/2, EXP.II –1/2). For the grammatical argument comparison *ARG*, only the Akan data were of interest, and thus sum contrasts were used to compare subject focus to object focus (a within-participant but between-item manipulation) with the weights for German set to zero (AK.SBJ 1/2, AK.OBJ –1/2, GE.SBJ 0). For the random-effects structures, varying intercepts as well as varying slopes for both the within-participant and within-item manipulations were used. Moreover, regularizing priors were used in order to downweight extreme values and obtain stable inferences (Vasishth et al. 2018).[27] See Appendix D.1 for the contrast coding for all main effects and interactions as well as the specification of the model reported here.

From the model I will report the posterior distribution of the parameters of interest and the 95% credible intervals (CrI)—i.e., given the data and model, the interval containing the 95% most credible values of the parameter. In cases when the credible interval of the estimates overlaps with zero, this will be taken as evidence of uncertainty about the effect, since it is possible that the effect is either zero (no effect) or has the wrong sign (e.g., the effect is in fact positive despite the posterior mean having a negative value). By contrast, when the credible interval does not overlap with zero at all, this will be interpreted as evidence of a robust effect. In cases when overlap with zero is minimal, with most of the probability density on one side of zero, this will be interpreted as weak evidence of an effect.

As can be seen in Figure 4.4, given that all of the credible intervals cross zero, no robust effect was found in any of the main comparisons. Moreover, for all estimates—especially for *EXP* and *LANG* as well as the interaction of the two factors—there was high uncertainty about the coefficient, seen in the broad credible intervals (logit scale). That no effect was found is perhaps unsurprising given the low power of the study and the potentially small effect for some comparisons (e.g., *ARG*). There was, however, a weak and unreliable effect for the three-way interaction of *ST.EXP.LANG* ($\hat{\beta} = 0.47$, CrI: [–0.08, 1.06]), for which I have no explanation at the time of writing. See Appendix D.1 for the model coefficients for all effects in both logit and probability scale, which for the sake of space I have not reported here.

27 Note that I am not computing Bayes factors here, which are highly sensitive to the prior, and in particular the prior uncertainty (Schad, Betancourt & Vasishth 2019).

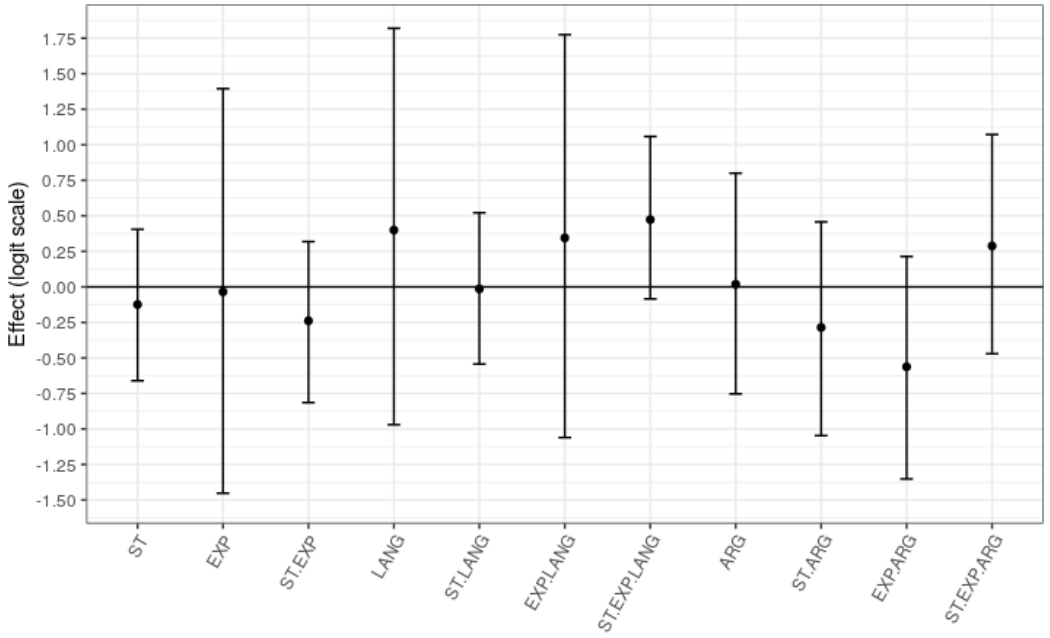


Figure 4.4 *Model estimates (posterior mean and 95% credible intervals) on the logit scale.*

4.3.2.2 Clefts and definite pseudoclefts: by participant results

Clefts and definite pseudoclefts in both Akan and German elicited a midway response on average. In De Veaugh-Geiss et al. (2018) it was reported that the intermediate response pattern in German was not due to individual participants responding at (near) chance levels with a normal distribution around the mean; rather, participants fell into two main clusters, i.e., those who largely judged the sentences as 'true' (*non-exhaustive* responders) and those who largely judged the sentences as 'false' (*exhaustive* responders). Moreover, a positive correlation between clefts and pseudoclefts was found: when participants judged clefts exhaustively they also judged definite pseudoclefts exhaustively, and vice versa for when participants judged clefts/definite pseudoclefts non-exhaustively.

In Akan the pattern is similar, as can be seen in the two plots in Figure 4.5. I will start with the left plot. In this plot one can see the counts of participants given the number of 'true' judgments made, which is split into two grids, one for clefts and one for pseudoclefts. Each participant had the possibility of making up to four judgments in the critical manipulations: when participants made 0/4 judgments, they are counted up in the leftmost bar per grid, and so on when participants made 1/4 judgments, 2/4 judgments, etc. Plotted in this way, it is appears that the by-participant

responses exhibit a bimodal distribution. In other words, for both sentence types in both languages the majority of participants either made 0/4 'true' judgments (exhaustive responders) or 4/4 'true' judgments (non-exhaustive responders), with fewer participants in between (although the results for definite pseudoclefts in Akan are not so clear-cut).

What one cannot see in the left plot is whether the same participants that judged clefts as exhaustive also judged definite pseudoclefts as exhaustive, and vice versa. This is shown in the right plot in Figure 4.5. In this plot, the x-axis shows the counts for clefts and the y-axis the counts for definite pseudoclefts, going from 0 to 4, and each point represents an individual participant. Akan participants are represented by circles and German participants by triangles; note that the symbols have been jittered around the exact values so the individual data points are not perfectly overlapping. The circles and triangles in the upper right corner show the participants who judged clefts and definite pseudoclefts as 'true' 4/4 times (non-exhaustive responders), while those in the bottom left corner show the participants who judged the two sentence types as 'true' 0/4 times (exhaustive responders). The two lines show the positive correlation between the two sentence types. Not a single participant is found in the upper-left or lower-right corner: these would be cases when a participant always judged one of the two sentence types exhaustively but the other non-exhaustively.

4.4 Discussion

It is critical to note right off the bat that the absence of evidence is not equal to evidence of absence—yet the most relevant findings to report here are the null effects in the various comparisons of interest. Moreover, in the participant random effects structure, at least one model assumption appears to be violated given the bimodal distribution of the by-participant responses. Nevertheless, in the discussion I will focus on the null effects and, in particular, the by-participant data for clefts and definite pseudoclefts, since this participant clustering toward the exhaustive or non-exhaustive responses is the primary motivation for the discourse pragmatic proposal in De Veaugh-Geiss et al. 2018.

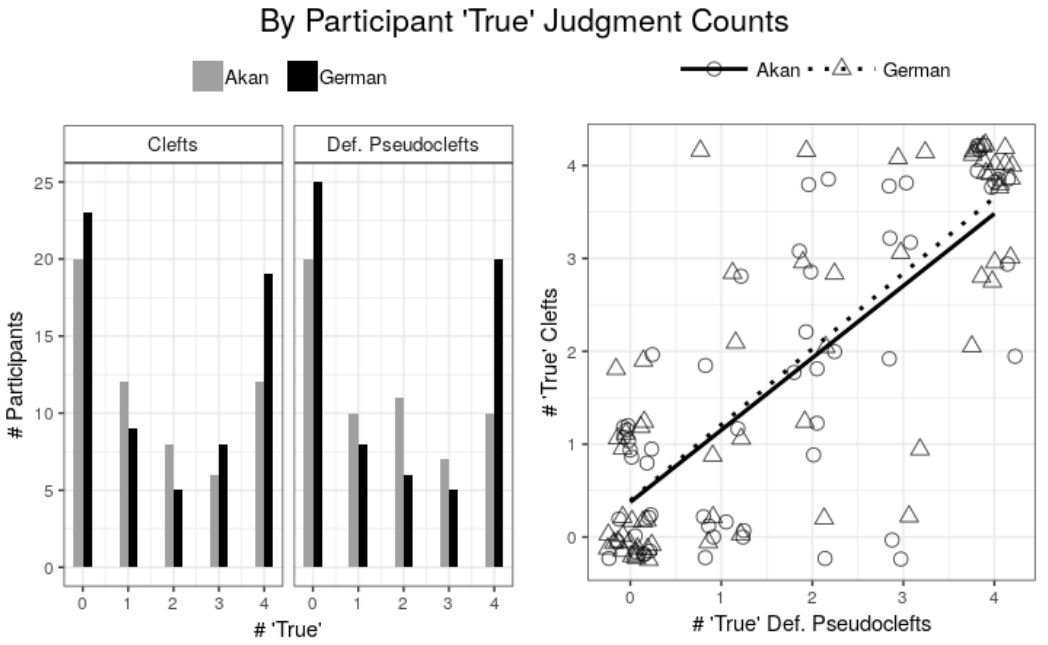


Figure 4.5 *'True' responses for the 122 participants in the cleft and definite pseudocleft conditions for the critical combinations* [+VER, –EXH] *and* [+FAL, +CAN]. *Left: counts of participants per total number of 'true' judgments made (out of 4 total per participant); a majority of participants either made 0/4 'true' judgments (exhaustive responders) or 4/4 'true' judgments (non-exhaustive responders), resulting in a bimodal distribution. Right: each circle (Akan) and triangle (German) represents an individual participant's counts of 'true' judgments for definite pseudoclefts (x-axis) and clefts (y-axis), with the lines showing a positive correlation.*

The three research questions, repeated here in brief for the reader, were as follows:

Q1 *Cross-linguistic variability*: Are *nà*-clefts and definite pseudoclefts of the form *Nipa no a* ... 'The person who ...' strongly exhaustive in Akan, in comparison to their counterparts in German?

Q2 *Inter-argument/inter-sentential variability*: Are Akan object clefts more strongly exhaustive than (i) subject clefts as well as (ii) definite pseudoclefts (for both grammatical arguments)?

Q3 *Inter-speaker variability*: Do participants cluster into 'exhaustivity' groups as was found in German?

For the first research question **Q1**, I entertained the hypothesis that one language might semantically-code exhaustivity in clefts/pseudoclefts, whereas another might not, which has been proposed for other languages and phenomena. For instance, recall that the SEMANTIC DEFINITE accounts predict that cleft and pseudocleft exhaustivity will be robust and systematic; in other words, exhaustivity in both sentence types should arise across contexts, speakers, and experimental settings. This is not what was found for either language. In fact—and in answer to the third research question **Q3**—Akan showed parallels to the bimodal response pattern reported for German. The DISCOURSE PRAGMATIC account, on the other hand, predicts that the exhaustivity inference in clefts and pseudoclefts can be variable, with either an exhaustive or non-exhaustive interpretation depending on the resolution of the anaphoric existence presupposition. Indeed, it is the latter approach which is compatible with the results reported for clefts and definite pseudoclefts across the two languages here. For the second research question **Q2**, I hypothesized that clefting for objects in Akan could give rise to a manner implicature strengthening the exhaustivity effect. However, this was not borne out in the data, as no statistical difference between subject and object grammatical arguments, nor an interaction of argument with sentence type, was found. In short, for the results of the experiments reported here, there is no reason to believe that exhaustivity in clefts and definite pseudoclefts differs either (i) across German and Akan or (ii) across grammatical arguments and sentences within Akan.

The discourse pragmatic approach in De Veaugh-Geiss et al. 2018 (following the dynamic account in Pollard & Yasavul 2016) takes clefts and definite pseudoclefts to be either strongly exhaustive or non-exhaustive depending on the resolution of the existence presupposition. (This approach is also the basis for the proposal in Destruel & De Veaugh-Geiss 2018, albeit with additional language-specific constraints placed by question-answer congruence.) The principle idea is tied to the observation that clefts come with an anaphoric existence presupposition (Delin 1992), which must

either (i) be bound when an antecedent is present or (ii) accommodated in the absence of an antecedent. De Veaugh-Geiss et al. (2018) exploit the anaphoricity of both clefts and definite pseudoclefts to model the parallel behavior across the two sentence types.

I will first discuss the existence presupposition before turning to anaphoricity. In Section 4.1 I stated that *nà*-clefts come with an existence presupposition, just as their German and English counterparts. This was shown for (1), repeated for the reader in (41).

(41) Kodwo na ɔ-di-i adua no.
Kodwo FM 3.SG.SBJ.ANI-eat-PST beans DET
'*It was Kodwo who ate the beans.*'
[adapted from ex. (6) in Pfeil, Genzel & Kügler 2015]

→ Someone ate the beans. (existential: *presupposed*)

However, cross-linguistically not all clefts are argued to have an existential encoded in the cleft structure. For instance, it has been argued that St'át'imcets and Nłeʔkepmxcin do not encode existence as part of the meaning of clefts in those languages (Davis, Matthewson & Shank 2004, von Fintel & Matthewson 2008, Koch & Zimmermann 2010). Similarly, Grubic, Renans & Duah (2019) provide evidence that the West-Chadic language Ngamo carries no existence presupposition. By contrast, Grubic, Renans & Duah (2019: §6) claim that clefts in Akan (as well as Ga, a Kwa language spoken in the Greater Accra Region of Ghana) do conventionally-code existence, which they argue is a presupposition, as has been argued for English and German, among other languages.[28]

To support the claim that clefts in Akan encode an existence presupposition, Grubic, Renans & Duah (2019) demonstrate that *nà*-clefts are unacceptable with a focused negative quantifier, as in (42b); compare to the acceptability of the *in situ* strategy in (42a), which does not encode an existential (cf. Geurts & van der Sandt 2004 for English focus). The principal idea behind this test is that there is a clash between the existence presupposition that there is an α such that he invited α and the assertion that there is no α that he invited.

(42) *Context*: Who did Kofi invite to the party?
a. ɔ-à-m̀-frɛ́ ɔ̀bíárá.
3.SG.SBJ-PST-NEG-invite everybody/anybody

28 Given their lack of a conventionally-coded existential it is perhaps not surprising that clefts in St'át'imcets, Nłeʔkepmxcin, and Ngamo are also reported to lack a strong exhaustivity effect, which is compatible with the discourse pragmatic approach proposed here. Exploring this correlation further is of high interest for future research.

'*He did not invite anybody.*' (*SVO*)

b. #è-ǹ-yɛ́ òbíárá nà ɔ-frɛ́-èɛ́.
3.SG.INA.SBJ-NEG-be everybody/anyone FM 3.SG.SBJ-invite-PST
'*It was NOBODY that he invited.*' (*cleft*)
[ex. (60) in Grubic, Renans & Duah 2019]

For definite pseudoclefts, as in (43), results are entirely parallel: the sentences were judged as odd by two native-speaker consultants.

(43) *Context*: Who did Kofi invite to the party?

a. #Nipa no a ɔ-frɛ́-ɛ́ ǹ-yɛ́ òbíárá.
person the who 3.SG.SBJ-invite-PST NEG-be anybody (*def. pse.*)
[adapted from ex. (60) in Grubic, Renans & Duah 2019]

Thus, I will follow Grubic, Renans & Duah (2019) and take Akan clefts and definite pseudoclefts to give rise to an existential presupposition.

This leads me to anaphoricity in the two sentence types, illustrated in (44). The two sentence types in (44a)–(44b) presuppose an existential, i.e., that there exists a person α with property *P*. The indefinite *òbí* in the previous sentence in (44) introduces a discourse referent α with the same property *P*, and thus the existence presupposition will be dynamically bound by this discourse antecedent. Indeed, an Akan consultant suggests that the anaphoric interpretation is the only possible reading of (44a)–(44b).

(44) Òbí tɔ̀-ɔ̀ àtàádé ńnórà. ...
someone buy-PST dress yesterday
'*Someone bought a dress yesterday. ...*'

a. Kumi na ɔ-tɔ́-ɔ̀ àtàádé.
Kumi FM 3.SG.SBJ-buy-PST dress
'*It is Kumi who bought a dress.*' (*cleft*)

b. Nipa no a ɔ-tɔ́-ɔ̀ àtàádé ne Kumi.
person the who 3.SG.SBJ-buy-PST dress be Kumi
'*The person who bought a dress is Kumi.*' (*def. pse.*)
[adapted from ex. (34) above]

The above are examples for when the existence presupposition is bound by an overt discourse antecedent; however, in the two experiments here the stimuli were presented out-of-the-blue and thus accommodation was necessary. In fact, the crux of the proposal in De Veaugh-Geiss et al. 2018 is in how participants accommodate the discourse antecedent—the strategy used will result in either a strongly exhaustive or a non-exhaustive interpretation. One option is that the existence pre-

supposition is resolved to the maximal discourse referent introduced by the implicit Question Under Discussion (QUD) *Who P?* (Roberts 1996, 2012). Following Hamblin (1973), questions are assumed to denote sets of true and complete answers, and the cleft/pseudocleft will pick up the maximal discourse referent introduced by this implicit QUD, resulting in an exhaustive interpretation. By contrast, for participants with a non-exhaustive interpretation, the existence presupposition is resolved to a non-maximal discourse referent which the cleft/pseudocleft specifies further, as for the indefinite antecedent in example (44).[29] Therefore, depending on how participants accommodate the anaphoric existence presupposition, a strongly exhaustive or non-exhaustive interpretation will be obtained. The proposal in De Veaugh-Geiss et al. 2018 can thus be extended to Akan to account for the fact that some participants got a non-exhaustive interpretation, while others got a strongly exhaustive interpretation, and this holds for both clefts and definite pseudoclefts.

4.5 Conclusion

In this paper I have presented two experiments on the exhaustivity inference in *nà*-clefts and definite pseudoclefts of the form '*Nipa no a* P *ne* α' ('The person who did *P* is α') in Akan, with a direct comparison to their counterparts in German using the same experimental set-up. The analysis presented here was an exploratory one given the unexpected uncover-first-judge-later response strategy employed by Akan participants. Nevertheless, despite the unforeseen response patterns in the incremental information-retrieval paradigm for Akan, the results of the post hoc analysis are compatible with a parallel approach both (i) cross-sententially between clefts and definite pseudoclefts, and (ii) cross-linguistically between Akan and German. Moreover, although Akan exhibits to some extent a subject/non-subject asymmetry for *ex situ* focus marking with *nà*-clefts, no difference was found in the strength of exhaustivity between the two grammatical arguments. Finally, the by-participant variability found in both languages can be accounted for with the discourse-pragmatic analysis of cleft exhaustivity from De Veaugh-Geiss et al. (2018) (based on the dynamic account of Pollard & Yasavul 2016), which has been adopted for focus-background clefts in various languages (De Veaugh-Geiss et al. 2018 for German; Zimmermann et al. to appear for English and Hungarian; and Destruel & De Veaugh-Geiss 2018, 2019 for French).

29 Onea (2013: p. 13) argues that discourse referents introduced by indefinites trigger a particular type of question called a potential question, informally described as a question that arises when new information enters a discourse and "interlocutors may get interested in particular aspects of that piece of information and they may [...] answer questions unrelated (or just loosely related) to the previous topic" (cf. questions and sub-questions in Roberts 1996; see also Onea 2016).

Bibliography

Abakah, Emmanuel N. 2005. Tone rules in Akan. *Journal of West African Languages* XXXII(1–2). 109–134 [p. 131].

Abbott, Barbara. 2014. The indefiniteness of definiteness. In Thomas Gamerschlag, Doris Gerland, Rainer Osswald & Wiebke Petersen (eds.), *Concept types and frames*. New York: Springer [pp. 12, 33].

Aboh, Enoch O. 2010. The morphosyntax of the noun phrase. In James Essegbey & Enoch O. Aboh (eds.), *Topics in Kwa syntax*, 11–38. Dordrecht: Springer [p. 131].

Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics and Philosophy* 34. 491–535 [p. 12].

Abrusán, Márta. 2016. Presupposition cancellation: Explaining the 'soft-hard' trigger distinction. *Natural Language Semantics* 24. 165–202. https://doi.org/10.1007/s11050-016-9122-7 [pp. 11, 12, 14, 15, 23, 69, 88, 89].

Abrusán, Márta & Kriszta Szendrői. 2013. Experimenting with the king of France: Topics, verifiability and definite descriptions. *Semantics and Pragmatics* 6(10). 1–43 [pp. 15, 16, 17, 18, 19, 37, 38, 85, 87].

Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. In Brendan Jackson (ed.), *Semantics and Linguistic Theory (SALT)*, vol. 12. Ithaca, NY: CLC Publications [pp. 12, 15, 89].

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27. Advance Access publication cited, 37–80. https://doi.org/10.1093/jos/ffp009 [pp. 14, 15, 89].

Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistics Inquiry* 1(2). 149–168 [pp. 7, 27, 145].

Altmann, Gerry & Mark Steedman. 1988. Interaction with context during sentence processing. *Cognition* 30. 191–238 [p. 74].

Altmann, Hans. 2009. Gradpartikel. In Ludger Hoffmann (ed.), *Handbuch der deutschen Wortarten*, 357–386. Walter de Gruyter. https://doi.org/10.1515/9783110217087.357 [p. 26].

Ameka, Felix K. 2010. Information packaging construction in Kwa: Microvariation and typology. In Essegbey James & E. O Aboh (eds.), *Topics in Kwa syntax*, 141–176. Dordrecht: Springer [pp. 128, 129, 133, 138].

Amfo, Nana Aba A. 2007. Noun phrase and clausal connectives in Akan. *Studies in African Linguistics* 36(1). 1–28 [p. 133].

Amfo, Nana Aba A. 2010. Lexical signaling of information structure in Akan. *Linguistics* 48(1). 195–225 [pp. 128, 129].

Arkoh, Ruby & Lisa Matthewson. 2013. A familiar definite article in Akan. *Lingua* 123. 1–30 [p. 135].

Atlas, Jay & Stephen Levinson. 1981. *It*-clefts, informativeness and logical form: Radical pragmatics. In Peter Cole (ed.), *Radical pragmatics*, 1–61. Academic Press [pp. 1, 69].

Audacity Team. 2018. *Audacity(R): Free Audio Editor And Recorder [Computer Application]*. Version 2.2.2. https://audacityteam.org/ [p. 114].

Bacovcin, Hezekiah Akiva, Jérémy Zehr & Florian Schwarz. 2018. To accommodate or to ignore? The presuppositions of 'again' and 'continue' across contexts. *Glossa: A Journal of General Linguistics* 3(1). 1–23. https://doi.org/10.5334/gjgl.402 [pp. 15, 19].

Bade, Nadine. 2015. *Obligatory presupposition triggers in discourse: Empirical investigations of the theories Maximize Presupposition and Obligatory Implicatures*. Universität Tübingen dissertation [p. 72].

Barker, Chris. 2004. Possessive weak definites. In Ji-yung Kim, Barbara Hall Partee & Yury A Lander (eds.), *Possessives and beyond: Semantics and syntax*, 89–113. MA: GLSA Publications [p. 33].

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. Parsimonious mixed models. Retrieved from http://arxiv.org/abs/1506.04967 (ArXiv e-print; submitted to *Journal of Memory and Language*) [pp. 47, 79].

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01 [p. 47].

Beaver, David & Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Wiley-Blackwell [pp. 8, 142].

Beaver, David & Edgar Onea. 2015. Cleftomania. Paper presented at the workshop *Pragmatics of Questions*, Amsterdam. Amsterdam [p. 28].

Beaver, David & Henk Zeevat. 2007. Accommodation. In Gilian Ramchand & Charles Reiss (eds.), *The Oxford handbook for lingusitic interfaces*, 502–538. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199247455.013.0017 [p. 121].

Berg, Janine, Marianne Furrer, Ellie Harmon, Uma Rani & M. Six Silberman. 2018. *Digital labour platforms and the future of work: Towards decent work in the*

online world. Geneva: International Labour Office, United Nations. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_645337.pdf [p. 118].

van Berkum, Jos J. A., Colin M. Brown, Peter Hagoort & Pienie Zwitserlood. 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology* 40. 235–248 [p. 74].

Beyssade, Claire, Barbara Hemforth, Jean-Marie Marandin & Cristel Portes. 2015. Prosodic realizations of information focus in French. In Lyn Frazier & Edward Gibson (eds.), *Explicit and implicit prosody in sentence processing*, 39–61. Amsterdam: John Benjamins Publishing Company [p. 77].

Boadi, Lawrence A. 1974. Focus marking in Akan. *Linguistics* 140. 5–57 [pp. 3, 7, 127, 128, 132, 133, 134, 135, 141, 143, 144, 145].

Boadi, Lawrence A. 2005. *Three major syntactic structures in Akan: Interrogatives, complementation, and relativisation*. Accra, Ghana: Black Mask Ltd. [p. 131].

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10). 341–345 [p. 114].

Boersma, Paul & David Weenink. 2018. *Praat: Doing Phonetics By Computer [Computer Application]*. Version 6.0.37. https://www.praat.org/ [p. 114].

Bombi, Carla. 2017. *Definiteness in Akan*. Author name in original: Carla Bombi Ferrer. Universität Potsdam MSc thesis [p. 131].

Bombi, Carla & Joseph P. De Veaugh-Geiss. 2018. Quantitative data in the field: Two case studies on Akan. Poster presented at *Linguistic Evidence* conference, 15 February, Universität Tübingen [pp. 131, 153, 155].

Bombi, Carla, Mira Grubic, Agata Renans & Reginald Akuoko Duah. 2019. The semantics of the clausal determiner *no* in Akan (Kwa). In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Sinn und Bedeutung (SuB)*, vol. 23, 181–199. Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès) [p. 135].

Bott, Lewis & Ira A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* (51). 437–457 [pp. 21, 73, 74, 86].

Breheny, Richard, Heather J. Ferguson & Napoleon Katsos. 2013. Investigating the time course of accessing conversational implicatures during incremental sentence interpretation. *Language and Cognitive Processes* 28. 443–467 [pp. 21, 73].

Breheny, Richard, Napoleon Katsos & John Williams. 2006. Are generalized scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100. 434–463 [pp. 21, 73, 74, 86].

Büring, Daniel. 2011. Conditional exhaustivity presuppositions in clefts (and definites). Ms. ZAS/Vienna [pp. 13, 142].

Büring, Daniel. 2015. Unalternative semantics. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 25, 550–575. https://doi.org/10.3765/salt.v25i0.3634 [p. 30].

Büring, Daniel & Manuel Križ. 2013. It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics* 6(6). 1–29. https://doi.org/10.3765/sp.6.6 [pp. 1, 3, 5, 6, 7, 9, 11, 13, 22, 25, 27, 31, 32, 33, 58, 66, 69, 93, 94, 96, 97, 98, 102, 120, 121, 122, 123, 129, 142, 145, 146].

Burkhardt, Petra. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language* 98. 159–168 [pp. 74, 87].

Bürkner, Paul-Christian. 2017. *brms*: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28. https://doi.org/10.18637/jss.v080.i01 [pp. 110, 163].

Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modeling with the R package *brms*. *The R Journal* 10(1). 395–411 [pp. 110, 163].

Bürkner, Paul-Christian & Matti Vuorre. 2018. Ordinal regression models in psychology: A tutorial. Unpublished; Last Updated: 2018-09-06 03:16 PM. https://osf.io/cu8jv [p. 110].

Byram-Washburn, Mary, Elsie Kaiser & Maria Luisa Zubizarreta. 2013. *The English* it-*Cleft: No Need To Get Exhausted*. Poster: Linguistic Society of America (LSA) [pp. 3, 29, 36, 71, 120, 129].

Cahill, Michael. 1985. *An autosegmental analysis of nasality and tone*. Texas: Arlington MA thesis [p. 131].

Carlson, Greg & Rachel Sussman. 2005. Seemingly indefinite definites. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives*, 71–86. https://doi.org/10.1515/9783110197549.71 [p. 33].

Carlson, Greg, Rachel Sussman, Natalie Klein & Michael K. Tanenhaus. 2006. Weak definite noun phrases. In Chris Davis, Amy Rose Deal & Youri Zabbal (eds.), *North Eastern Linguistic Society (NELS)*, vol. 36, 179–196. MA: GLSA Publications [p. 33].

Carter-Thomas, Shirley. 2009. The French *c'est*-cleft: Functional and formal motivations. In David Banks, Simon Eason & Janet Ormrod (eds.), *La linguistique systémique fontionnelle et la langue française*, 127–156. Paris, France: L'Harmattan [p. 68].

Champollion, Lucas. to appear. Distributivity, collectivity, and cumulativity. In Lisa Matthewson, Cécile Meier, Hotze Rullman & Thomas Ede Zimmermann (eds.), *Wiley's companion to semantics*. Wiley, Hoboken, NJ [p. 95].

Chemla, Emmanuel & Lewis Bott. 2013. Processing presuppositions: Dynamic semantics vs .pragmatic enrichment. *Language and Cognitive Processes* 38(3). 241–260 [p. 18].

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and beyond*, 39–103. Oxford University Press [p. 73].

Chierchia, Gennaro & Sally McConnell-Ginet. 1996. *Meaning and grammar: An introduction to semantics*. MIT Press [pp. 12, 13].

Clech-Darbon, Anne, Georges Rebuschi & Annie Rialland. 1999. Are there cleft sentences in French? In Georges Rebuschi & Laurice Tuller (eds.), *The grammar of focus*, 83–118. Amsterdam: John Benjamins Publishing Company [pp. 23, 69].

Collins, Peter C. 1991. *Cleft and pseudo-cleft constructions in English*. London: Routledge [p. 129].

Conroy, Anastasia Marie. 2008. *The role of verification strategies in semantic ambiguity resolution in children and adults*. University of Maryland dissertation [pp. 29, 151].

Coppock, Elizabeth & David Beaver. 2015. Definiteness and determinacy. *Linguistics and Philosophy* (38). 377–435. https://doi.org/10.1007/s10988-015-9178-8 [pp. 28, 60].

Davis, Henry, Lisa Matthewson & Scott Shank. 2004. Clefts vs. nominal predicates in two Salish languages. In Donna B. Gerdts & Lisa Matthewson (eds.), *Studies in Salish linguistics in honor of M. Dale Kinkade (University of Montana Occasional Papers in Linguistics 10)*, 100–117. Missoula: Montana: University of Montana Press [pp. 161, 169].

De Cesare, Anna-Maria & Davide Garassino. 2015. On the status of exhaustiveness in cleft sentences: An empirical and cross-linguistic study of English *also-*/*only*-clefts and Italian *anche-*/*solo*-clefts. *Folia Linguistica* 49. 1–56 [p. 129].

De Veaugh-Geiss, Joseph P. in preparation. *nà*-Cleft (non-)exhaustivity: Variability in Akan [pp. 3, 5, 11, 16, 20, 21, 23].

De Veaugh-Geiss, Joseph P., Swantje Tönnis, Edgar Onea & Malte Zimmermann. 2017. An experimental investigation of (non-)exhaustivity in *es*-clefts. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Sinn und Bedeutung (SuB)*, vol. 21, 347–365 [pp. 71, 151].

De Veaugh-Geiss, Joseph P., Swantje Tönnis, Edgar Onea & Malte Zimmermann. 2018. That's not quite it: An experimental investigation of (non-)exhaustivity in clefts. *Semantics and Pragmatics* 11(3). Early access. https://doi.org/10.3765/sp.11.3 [pp. 3, 5, 6, 7, 11, 12, 15, 16, 20, 21, 23, 65, 70, 71, 84, 88, 89, 90, 94, 98, 127, 128, 129, 130, 143, 145, 146, 147, 148, 149, 150, 151, 155, 156, 159, 160, 161, 165, 166, 168, 169, 170, 171].

De Veaugh-Geiss, Joseph P., Malte Zimmermann, Edgar Onea & Anna-Christina Boell. 2015. Contradicting (not-)at-issueness in exclusives and clefts: An empirical study. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 25, 373–393. First author was published under the name: DeVeaugh-Geiss, Joseph P. https://doi.org/10.3765/salt.v25i0.3054 [pp. 1, 23, 27, 29, 30, 36, 60, 69, 70, 71, 84, 87, 98, 129, 130, 134, 139, 146, 147].

DeCat, Cecile. 2007. *French dislocation: Interpretation, syntax and acquisition*. Oxford: Oxford University Press [p. 67].

Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Berlin: Mouton de Gruyter [p. 129].

Degen, Judith. 2015. Investigating the distribution of some (but not all) implicatures using corpora and web-based methods. *Semantics and Pragmatics* 8(11). 1–55. https://doi.org/10.3765/sp.8.11 [pp. 74, 86].

Degen, Judith & Michael K. Tanenhaus. 2016. Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive Science* 40. 172–201 [pp. 74, 86].

Delais-Roussarie, Elisabeth, Annie Rialland, Doetjes Jenny & Jean-Marie Marandin. 2002. The prosody of post-focus sequences in French. In Bernard Bel & Isabel Marlien (eds.), *Speech prosody*, 239–242. Aix en Provence [p. 77].

Delin, Judy. 1992. Properties of *it*-cleft presupposition. *Journal of Semantics* 9(4). 289–306. https://doi.org/10.1093/jos/9.4.289 [pp. 13, 27, 56, 70, 89, 146, 168].

Delin, Judy & Jon Oberlander. 1995. Syntactic constraints on discourse structure: The case of *it*-clefts. *Linguistics* 33. 456–500. https://doi.org/10.1515/ling.1995.33.3.465 [p. 2].

Delin, Judy & Jon Oberlander. 2005. Cleft constructions in context: Some suggestions for research methodology. MS., University of Stirling. http://www.fb10.uni-bremen.de/anglistik/langpro/projects/GeM/delin-publications.html [p. 129].

Destruel, Emilie. 2012. The French *c'est*-cleft: An empirical study on its meaning and use. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics (EISS)*, vol. 9, 95–112. Paris: CSSP [p. 29].

Destruel, Emilie. 2013. *An empirical study on the meaning and use of the French* c'est-*cleft*. University of Texas at Austin dissertation [pp. 68, 71].

Destruel, Emilie, David Beaver & Elizabeth Coppock. 2017. Clefts: Quite the contrary! In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Sinn und Bedeutung (SuB)*, vol. 21, 335–346 [pp. 36, 67, 120].

Destruel, Emilie & Joseph P. De Veaugh-Geiss. 2018. On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts. *Journal of Pragmatics* 138. Second author was published under the name: DeVeaugh-

Geiss, Joseph P., 1–16. https://doi.org/10.1016/j.pragma.2018.09.009 [pp. 3, 4, 5, 16, 20, 21, 23, 24, 98, 133, 146, 148, 168, 171].

Destruel, Emilie & Joseph P. De Veaugh-Geiss. 2019. (Non-)exhaustivity in French *c'est*-clefts. In Christopher Pinon (ed.), *Empirical Issues in Syntax and Semantics (EISS)*, vol. 12, 91–120. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ [pp. 1, 129, 130, 145, 146, 151, 171].

Destruel, Emilie, Daniel Velleman, Edgar Onea, Dylan Bumford, Jingyang Xue & David Beaver. 2015. A cross-linguistic study of the non-at-issueness of exhaustive inferences. In Florian Schwarz (ed.), *Experimental perspectives on presuppositions*, 135–156. Springer. https://doi.org/10.1007/978-3-319-07980-6_6 [pp. 3, 28, 29, 66, 68, 71, 90, 94, 98, 120, 129].

Destruel, Emilie & Leah Velleman. 2014. Refining contrast: Empirical evidence from the English *it*-cleft. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics (EISS)*, vol. 10, 197–214. Paris: CSSP [pp. 36, 67, 68, 90, 120].

Dik, Simon, Maria E. Hoffman, de Jong, Jan R., Sie Ing Djiang, Harry Stroomer & Lourens de Vries. 1981. On the typology of focus phenomenon. In Harry van der Hulst & Micheal Moortgat (eds.), *Perspectives on functional grammar*. Dordrecht: Foris [p. 129].

den Dikken, Marcel. 2013. Predication and specification in the syntax of cleft sentences. In Katharina Hartmann & Tonjes Veenstra (eds.), *Cleft structures*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.208.02dik [p. 2].

Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Ghana: University Press [p. 131].

Dolphyne, Florence Abena & Mary E. Kropp Dakubu. 1988. The Volta-Comoe languages. In M.E. Kropp Dakubu (ed.), *The languages of Ghana*, 50–90 [p. 131].

Donnellan, Keith. 1966. Reference and definite descriptions. *Philosophical Review* 75. 281–304 [p. 16].

Drenhaus, Heiner, Malte Zimmermann & Shravan Vasishth. 2011. Exhaustiveness effects in clefts are not truth-functional. *Journal of Neurolinguistics* 24. 320–337. https://doi.org/10.1016/j.jneuroling.2010.10.004 [pp. 66, 129].

Dryer, Matthew. 1996. Focus, pragmatic presupposition and activated propositions. *Journal of Pragmatics* 26(4). 473–523. https://doi.org/10.1016/0378-2166(95)00059-3 [p. 27].

Duah, Reginald Akuoko. 2013. *Force dynamics and causation in Akan*. University of Ghana dissertation. https://web.archive.org/web/20181204130737/http://ugspace.ug.edu.gh/bitstream/handle/123456789/5192/Reginald%5C%20Akuoko%5C%20Duah_Force%5C%20Dynamics%5C%20and%5C%20Causation%5C%20in%5C%20Akan_2013.pdf [p. 131].

Duah, Reginald Akuoko. 2015. Exhaustive focus marking in Akan. In F. Bildhauer & M. Grubic (eds.), *Interdisciplinary Studies on Information Structure (ISIS) 19:*

Mood, exhaustivity & focus marking in non-European languages, 1–28 [pp. 127, 128, 129, 133, 134, 135, 136, 137, 138, 139, 141, 142].

Duah, Reginald Akuoko. 2019. Coordination, tense, focus and the road in between: The case of the particle *na* in Akan. Handout: Talk given at the SynSem Colloquium, Universität Potsdam, 7 May 2019 [p. 133].

Dufter, Andreas. 2009. Clefting and discourse organization: Comparing Germanic and Romance. In Andreas Dufter & Daniel Jacob (eds.), *Focus and background in Romance languages (= studies in language companion series 112)*, 83–121. Amsterdam: John Benjamins Publishing Company [p. 90].

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273 [pp. 66, 68, 85].

É. Kiss, Katalin. 1999. The English cleft construction as a focus phrase. *Boundaries of Morphology and Syntax*. 217–229 [pp. 2, 26].

É. Kiss, Katalin. 2010. Structural focus and exhaustiveness. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological and experimental perspectives*, 64–88. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199570959.003.0004 [p. 55].

Ellis, Jeffrey & Lawrence A. Boadi. 1969. To 'be' in Twi. In John W.M. Verhaar (ed.), *The verb 'be' and its synonyms, part 4*, 1–71. Dordrecht: D. Reidel [p. 128].

Féry, Caroline. 2013. Focus as alignment. *Natural Language and Linguistic Theory* (31). 683–734 [pp. 67, 68].

Fiedler, Ines, Katharina Hartmann, Brigette Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological and experimental perspectives*, 234–257. Oxford: Oxford University Press [p. 138].

Fiedler, Ines & Anne Schwarz. 2005. Out-of-focus encoding in Gur and Kwa. In Shinichiro Ishihara, Michaela Schmitz & Anne Schwarz (eds.), *Interdisciplinary Studies on Information Structure (ISIS) 3*, 111–142 [p. 138].

Fiedler, Ines & Anne Schwarz. 2007. Narrative focus strategies in Gur and Kwa. In Aboh Enoch, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in Niger-Congo and Afroasiatic – on the interaction of focus and grammar in some African languages*, 267–286. Berlin: Mouton de Gruyter [pp. 135, 138].

von Fintel, Kai. 2003. Pragmatics: Notes on presupposition. MIT [p. 13].

von Fintel, Kai. 2008. What is a presupposition accommodation, again? *Philosophical Perspectives* 22.1. 137–170 [p. 121].

von Fintel, Kai & Lisa Matthewson. 2008. Universals in semantics. *The Linguistic Review* 25(1–2). 139–201 [p. 169].

Fodor, Janet Dean. 2002a. Prosodic disambiguation in silent reading. In Masako Hirotani (ed.), *North Eastern Linguistic Society (NELS)*, vol. 32, 113–132. Amherst, MA: GLSA [p. 113].

Fodor, Janet Dean. 2002b. Psycholinguistics cannot escape prosody. In *International Conference on Speech Prosody*. Aix en Provence, France [p. 113].

Fominyam, Henry & Radek Šimík. 2017. The morphosyntax of exhaustive focus: A view from Awing (Grassfields Bantu). *Natural Language and Linguistic Theory* 35. 1027–1077. https://doi.org/10.1007/s11049-017-9363-2 [p. 139].

Franke, Michael, Fabian Schlotterbeck & Petra Augurzky. 2016. Embedded scalars, preferred readings and prosody: An experimental revisit. *Journal of Semantics* 34(1). 153–199. https://doi.org/10.1093/jos/ffw007 [pp. 29, 151].

Fraurud, Kari. 1990. Definiteness and the processing of noun phrases in natural discourse. *Journal of Semantics* 7. 395–433 [p. 13].

Frazier, Lyn, Jeremy M. Pach & Keith Rayner. 1999. Taking on semantics commitments, II: Collective vs. distributive readings. *Cognition* 70(1). 87–104. https://doi.org/10.1016/s0010-0277(99)00002-5 [p. 96].

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*. English Translation: *On Sense and Meaning*. In Brian McGuinness (ed.), *Frege: Collected Works*, 157–177. Oxford: Basil Blackwell., 25–50 [p. 60].

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition and logical form*. New York: Academic Press [p. 15].

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Harvard University Press & Blackwell [pp. 2, 26].

Genzel, Susanne. 2013. *Lexical and post-lexical tones in Akan*. Potsdam, Germany: Universität Potsdam dissertation. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/7796/file/genzel_diss.pdf [pp. 131, 132, 138, 141, 148].

Gerőcs, Mátyás, Anna Babarczy & Balázs Surányi. 2014. Exhaustivity in focus: Experimental evidence from Hungarian. In Joseph Emonds & Markéta Janebová (eds.), *Language use and linguistic structure*, 181–194. Olomouc: Palacky University [pp. 71, 72].

Geurts, Bart & Rob van der Sandt. 2004. Interpreting focus. *Theoretical Linguistics* 30. 1–44 [p. 169].

Grice, Paul Herbert. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Speech acts (Syntax and Semantics 3)*, 41–58. New York: Academic Press [pp. 15, 70, 148].

Grice, Paul Herbert. 1981. Presupposition and implicature. In Peter Cole (ed.), *Radical pragmatics*, 183–199. Reprinted in Paul Herbert Grice, 1989, *Studies in the Way of Words*, Harvard University Press, Cambridge, Massachusetts, pp. 269–282. New York: Academic Press [p. 13].

Grodner, Daniel J., Natalie M. Klein, Kathleen M. Carbary & Michael K. Tanenhaus. 2010. ‘‘Some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116. 42–55 [pp. 21, 73, 86].

Grubic, Mira, Agata Renans & Reginald Akuoko Duah. 2019. Focus, exhaustivity and existence in Akan, Ga and Ngamo. *Linguistics* 57(1). 221–268 [pp. 1, 127, 128, 129, 132, 133, 134, 135, 137, 140, 142, 143, 160, 161, 169, 170].

Grubic, Mira & Marta Wierzba. 2019. Presupposition accommodation of the German additive particle *auch* (= “too”). *Frontiers in Communication* 4. 15. https://doi.org/10.3389/fcomm.2019.00015 [p. 19].

Guerts, Bart. 1998. The mechanism of denial. *Language* 74(2). 274–307. https://doi.org/10.1353/lan.1998.0264 [p. 121].

Halvorsen, Per-Kristian. 1978. *The syntax and semantics of cleft constructions*. University of Texas at Austin dissertation [p. 9].

Hamblin, Charles Leonard. 1957. *Language and the theory of information*. London School of Economics & Political Science (University of London) dissertation [p. 58].

Hamblin, Charles Leonard. 1971. Mathematical models of dialogue. *Theoria* 2. 131–155 [p. 70].

Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53 [pp. 89, 171].

Hamlaoui, Fatima. 2009. *La focalisation à l’interface de la syntaxe et de la phonologie: Le cas du français dans une perspective typologique*. Paris 3 Sorbonne Nouvelle dissertation [p. 68].

Hamlaoui, Fatima, Sascha Coridun & Caroline Féry. 2012. Expression prosodique du focus et du donné au sein des groupes nominaux [N A] du français. In F. Neveu, V. Muni Toke, P. Blumenthal, T. Klinger, P. Ligas, S. Prévost & S. Teson-Bonnard (eds.), *Actes du 3e Congrès Mondial de Linguistique Française*, 1505–1518. EDP Sciences [p. 77].

Han, Chung-Hye & Nancy Hedberg. 2008. The syntax and semantics of *it*-clefts: A tree-adjoining grammar analysis. *Journal of Semantics* 25(4). 345–380. https://doi.org/10.1093/jos/ffn007 [p. 7].

Harries-Delisle, Helga. 1978. Contrastive emphasis and cleft sentences. In Joseph H. Greenberg (ed.), *Universals of human language: Syntax*, 419–486. Stanford, CA: Stanford University Press [p. 7].

Hartmann, Katharina & Malte Zimmermann. 2007. In place, out of place? Focus in Hausa. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 365–403. Amsterdam: John Benjamins [p. 132].

Hedberg, Nancy. 1990. *Discourse pragmatics and cleft sentences in English*. University of Minnesota dissertation [p. 2].

Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76(4). 891–920. https://doi.org/10.2307/417203 [pp. 7, 27, 56].

Hedberg, Nancy. 2013. Multiple focus and cleft sentences. In Katharina Hartmann & Tonjes Veenstra (eds.), *Cleft structures*, 227–250. Amsterdam: John Benjamins. https://doi.org/10.1075/la.208 [pp. 26, 69].

Heim, Irene. 1982. *On the semantics of definite and indefinite noun phrases*. University of Massachusetts at Amherst dissertation [pp. 33, 60].

Heim, Irene. 1983. On the projection problem for presuppositions. In Michael Barlow, Daniel Flickinger & N. Wiegand (eds.), *2nd West Coast Conference on Formal Linguistics*, 114–126. Reprinted in Portner P., Partee B., eds. Formal Semantics: The Essential Readings. Oxford: Blackwell; 2002, 249–260. Stanford, California: Stanford University Press [p. 15].

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9. 183–221 [p. 12].

Heim, Irene. 1994. Plurals. Ms. Massachusetts Institute of Technology [p. 96].

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell [pp. 12, 14, 98].

Heizmann, Tanja. 2007. Children's acquisition of exhaustivity in clefts. In Heather Caunt-Nulton, Samantha Kulatilake & I-Hao Woo (eds.), *31st Boston University Conference on Language Development*, 298–309. Somerville, MA: Cascadilla Press [p. 72].

Heizmann, Tanja. 2012. *Exhaustivity in questions & clefts and the quantifier connection: A study in German and English*. University of Massachusetts, MA dissertation [p. 72].

Hemforth, Barbara. 2018. Linguistic evidence across languages. Invited talk at *Linguistic Evidence* conference, Universität Tübingen [p. 112].

Hole, Daniel & Malte Zimmermann. 2013. Syntactic partitioning in (South) East Asian: A cross-linguistic comparison of clefting in Japanese, Burmese and Chinese. In Katharina Hartmann & Tonjes Veenstra (eds.), *Cleft structures*, 285–318. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/la.208 [p. 71].

Horn, Laurence R. 1981. Exhaustiveness and the semantics of clefts. In Victoria E. Burke & James Pustejovsky (eds.), *North Eastern Linguistic Society (NELS)*, vol. 11, 125–142. Amherst: University of Massachusetts [pp. 1, 5, 9, 11, 23, 27, 30, 31, 32, 35, 56, 58, 60, 66, 69, 70, 84, 88, 93, 94, 98, 121, 129, 146, 147, 148].

Horn, Laurence R. 1985. Metalinguistic negation and pragmatic ambiguity. *Language* 61. 121–174. https://doi.org/10.2307/413423 [pp. 94, 121].

Horn, Laurence R. 1989. *A natural history of negation*. Reissued with new introduction (2001), David Hume Series, CSLI Publications, Stanford. University of Chicago Press [pp. 15, 94, 121, 122, 123].

Horn, Laurence R. 2014. Information structure and the landscape of (non-)at-issue meaning. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 108–127. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642670.013.009 [pp. 23, 27, 31, 35, 36, 60, 69, 70, 84, 129, 146, 147].

Horn, Laurence R. & Barbara Abbott. 2016. <the, a>: (In)definiteness and implicature. In Joseph Keim Campbell, William Kabasenche & Michael O'Rourke (eds.), *Reference and referring (topics in contemporary philosophy)*, vol. 10, 325–356. Cambridge, MA: MIT Press [p. 32].

Huang, Yi Ting & Jesse Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology* 58(3). 376–415 [pp. 21, 73, 74].

Huber, Stefan. 2002. *Es-Clefts und det-Clefts: Zur Syntax, Semantik und Informationsstruktur von Spaltsätzen im Deutschen und Schwedischen*. Stockholm: Almqvist & Wiksell International dissertation [p. 26].

Jespersen, Otto. 1927. *A modern English grammar on historical principles. Part III: Syntax. Second volume*. Heidelberg: Carl Winters [p. 2].

Jun, Sun-Ah & Cécile Fougeron. 2000. A phonological model of French intonation. In Antonis Botinis (ed.), *Intonation: Analysis, modeling and technology*, 209–242. Dordrecht: Kluwer [p. 77].

Kadmon, Nirit. 2001. *Formal pragmatics: Semantics, pragmatics, presupposition, and focus*. Malden: Blackwell [p. 13].

Karssenberg, Lena & Karen Lahousse. 2018. The information structure of French *il y a* clefts and *c'est* clefts: A corpus-based analysis. *Linguistics* 56. 513–548 [p. 90].

Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4. 55–69. https://doi.org/10.1080/08351817109370248 [p. 52].

Karttunen, Lauri. 1973. Presuppositions of compound sentences. *Linguistic Inquiry* 4. 167–193 [p. 12].

Karttunen, Lauri. 1974. Presuppositions and linguistic context. *Theoretical Linguistics* 1. 181–194 [pp. 8, 13].

Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyu Oh & David A. Dineen (eds.), *Presupposition* (Syntax and Semantics 11), 1–56. New York: Academic Press [p. 121].

Katz Bourns, Stacey. 2014. Contrasting *c'est*-clefts and *it*-clefts in discourse. In Stacey Katz Bourns & Lindsey L. Myers (eds.), *Perspectives on linguistic structure and context: Studies in honor of Knud Lambrecht*, 199–222. Amsterdam: John Benjamins Publishing Company [p. 68].

Katz, Stacey. 1997. *The syntactic and pragmatic properties of the* c'est-*cleft construction*. Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT 9802920). University of Texas at Austin dissertation [p. 67].

Katz, Stacey. 2000. Categories of *c'est*-cleft constructions. *Canadian Journal of Linguistics*. 253–273 [pp. 23, 69, 133].

Kobele, Gregory & Harold Torrence. 2006. Intervention and focus in Asante Twi. *ZAS Papers in Linguistics* (46) [pp. 128, 131, 133, 134, 135, 136, 137].

Koch, Karsten & Malte Zimmermann. 2010. Focus-sensitive operators in nłe?kepmxcin (Thompson River Salish). In Martin Prinzhorn, Viola Schmitt & Sarah Zobel (eds.), *Sinn und Bedeutung (SuB)*, vol. 14, 237–255. Vienna: Universität Wien [pp. 161, 169].

Koizumi, Yukiko. 2009. *Processing the 'not-because' ambiguity in English: The role of pragmatics and prosody*. City University of New York dissertation [p. 113].

Kratzer, Angelika. 2008. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Language context and cognition*, 269–299. Walter de Gruyter [p. 96].

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276 [pp. 123, 132].

Križ, Manuel. 2015. *it*-clefts, exhaustivity, and definite descriptions. In Klaus von Heusinger, Edgar Onea & Malte Zimmermann (eds.), *Questions in discourse*. To appear. Amsterdam: Brill [p. 28].

Križ, Manuel. 2016. Homogeneity, non-maximality, and 'all'. *Journal of Semantics* 33(3). https://doi.org/10.1093/jos/ffv006 [pp. 3, 6, 7, 22, 93, 94, 125, 145, 146].

Križ, Manuel. 2017. Referentiality, exhaustivity, and trivalence in *it*-clefts. Manuscript, Institut Jean-Nicod. The version cited here is archived at https://web.archive.org/web/20181123111403/https://semanticsarchive.net/Archive/TVjNDU0M/Kriz%202017%20Clefts.pdf [pp. 2, 3, 6, 7, 10, 22, 32, 33, 93, 94, 96, 97, 102, 103, 114, 121, 122, 123, 124, 134, 145, 146].

Križ, Manuel & Emmanuel Chemla. 2015. Two methods to find truth-value gaps and their application to the projection problem of homogeneity. *Natural Language Semantics* 23. 205–248. https://doi.org/10.1007/s11050-015-9114-z [p. 15].

Kruschke, John K. & Torrin M. Liddell. 2018. Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review* 25(1). 155–177. https://doi.org/10.3758/s13423-017-1272-1 [pp. 110, 163].

Kügler, Frank & Susanne Genzel. 2012. On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan. *Language and Speech* 55(3). 331–59. https://doi.org/10.1177/0023830911422182 [p. 133].

Kuppevelt, Jan. 1996. Inferring from topics. *Linguistics and Philosophy* 19(4). 393–443. https://doi.org/10.1007/bf00630897 [p. 15].

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and mental representations of discourse referents*. Cambridge University Press [pp. 66, 67, 68, 82].

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39. 463–516 [pp. 23, 69, 133].

Landman, Fred. 1989. Groups I. *Linguistics and Philosophy* 12(5). 559–605. https://doi.org/10.1007/BF00627774 [p. 95].

Langendoen, D. Terence & Harris Savin. 1971. The projection problem for presuppositions. In Charles Fillmore & D. Terence Langendoen (eds.), *Studies in linguistic semantics*, 373–388. Reprinted in János Petöfi and Dorothea Franck (eds) 1973, *Präsuppositionen in Philosophie und Linguistik — Presuppositions in Philosophy and Linguistics*, Athanäum Verlag, Frankfurt. New York: Holt, Reinhardt & Winston [p. 13].

Lasersohn, Peter. 1998. Generalized distributivity operators. *Linguistics and Philosophy* 21(1). 83–93. https://doi.org/10.1023/A:1005317815339 [p. 117].

Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press [p. 13].

Levinson, Stephen. 2000. *Presumptive meanings*. Cambridge: MIT Press [pp. 70, 73].

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8. Also in Rainer Bäuerle, Urs Egli and Arnim von Stechow (eds) 1979, *Semantics from Different Points of View*, Springer, Berlin, 339–359 [p. 13].

Lewis, Paul M., Gary F. Simons & Charles D. Fennig. 2013. *Ethnologue: Languages of the world*. Seventeenth. Dallas, Texas: SIL International. http://www.ethnologue.com [p. 131].

Liddell, Torrin M. & John K. Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. https://doi.org/10.1016/j.jesp.2018.08.009 [p. 110].

Ludlow, Peter & Gabriel Segal. 2004. On a unitary semantical analysis for definite and indefinite descriptions. In Anne Bezuidenhout & Marga Reimer (eds.), *Descriptions and beyond: An interdisciplinary collection of essays on definite and indefinite descriptions*. Oxford: Oxford University Press [pp. 12, 28].

Marfo, Charles & Adams Bodomo. 2005. Information structuring in Akan question-word fronting and focus constructions. *Studies in African Linguistics* 34(2). 179–208 [pp. 133, 137, 138].

Matthewson, Lisa. 2008. Pronouns, presuppositions, and semantic variation. In Tova Friedman & Satoshi Ito (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 18, 527–558 [pp. 85, 161].

Max Planck Institute for Evolutionary Anthropology, Department of Linguistics. 2018. *Leipzig Glossing Rules*. Last change: 31 May 2015. Cited version archived

at: https://web.archive.org/web/20181213114149/https://www.eva.mpg.de/lingua/resources/glossing-rules [p. 128].

Mayol, Laia & Elena Castroviejo. 2013. How to cancel an implicature. *Journal of Pragmatics* 50. 84–104. https://doi.org/10.1016/j.pragma.2013.02.002 [pp. 15, 35].

Moltman, Friederike. 1997. *Parts and wholes in semantics*. Oxford Univeristy Press [p. 96].

Moltman, Friederike. 2005. Part structures in situations: The semantics of *Individual* and *Whole*. *Linguistics and Philosophy* 28(5). 599–641. https://doi.org/10.1007/s10988-005-2142-2. [p. 96].

Moulton, Keir. 2007. Small antecedents: Syntax or pragmatics? In Emily Elfner & Martin Walkow (eds.), *North Eastern Linguistic Society (NELS)*, vol. 37. Amherst, MA: GLSA Publications [p. 15].

Nieuwland, Mante S., Tali Ditman & Gina R. Kuperberg. 2010. On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language* (63). 324–346 [p. 73].

Noveck, Ira A. & Andres Posada. 2003. Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language* 85. 203–210 [p. 86].

Ofori, Seth Antwi. 2011. On the basic focus marker, and the basic focus sentence, in Akan (Twi). *Nordic Journal of African Studies* 20(3). 241–262 [pp. 7, 127, 128, 132, 133, 134, 135, 144, 145].

Onea, Edgar. 2007. Exhaustivity, focus and incorporation in Hungarian. In Maria Aloni, Paul Dekker & Floris Roelofsen (eds.), *Sixteenth Amsterdam Colloquium*, 169–174. Author was published under the name: Onea Gáspár, Edgar. Amsterdam [p. 107].

Onea, Edgar. 2013. *Potential questions in discourse and grammar*. Göttingen: Georg-August-Universität Göttingen habilitation thesis [p. 171].

Onea, Edgar. 2016. *Potential questions at the semantics-pragmatics interface*. Leiden: Brill [pp. 58, 59, 90, 171].

Onea, Edgar & David Beaver. 2009. Hungarian focus is not exhausted. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 19, 342–359. https://doi.org/10.3765/salt.v19i0.2524 [pp. 3, 29, 35, 36, 71, 129].

Onea, Edgar & David Beaver. 2016. Complete vs. exhaustive answers: The discourse semantics of *it*-clefts. Talk given in Frankfurt on 24 May 2016 [p. 90].

Osam, E. Kweku. 2004. *Trondheim lectures – an introduction to the structure of Akan: Its verbal and multiverbal systems*. Legon: Department of Linguistics, University of Ghana [p. 131].

Patten, Amanda. 2012. *The English* it*-cleft: A constructional account and a diachronic investigation*. Berlin: De Gruyter Mouton [p. 129].

Percus, Orin. 1997. Prying open the cleft. In Kiyomi Kusumoto (ed.), *North Eastern Linguistic Society (NELS)*, vol. 27, 337–351 [pp. 1, 2, 3, 5, 7, 8, 22, 25, 27, 32, 33, 57, 60, 69, 94, 142, 145, 146].

Pfeil, Simone, Susanne Genzel & Frank Kügler. 2015. Empirical investigation of focus and exhaustivity in Akan. In Felix Bildhauer & Mira Grubic (eds.), *Interdisciplinary Studies on Information Structure (ISIS) 19: Mood, exhaustivity & focus marking in non-European languages*, 87–109 [pp. 128, 129, 139, 141, 169].

Pintér, Lilla. 2015. The acquisition of the exhaustive interpretation of Hungarian focus constructions. Talk Presented at the 12th International Conference on the Structure of Hungarian. Leiden: Germany [p. 72].

Pintér, Lilla. 2016. Preschoolers' interpretation of presupposed and pragmatically implied exhaustivity. Talk presented at the VIII International Conference of Language Acquisition. Palma de Mallorca: Spain [p. 72].

Poesio, Massimo & Renata Viera. 1998. A corpus-based investigation of definite description use. *Computational Linguistics* 24(2). 183–216 [p. 13].

Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. CSLI Publishing [pp. 2, 26].

Pollard, Carl & Murat Yasavul. 2016. Anaphoric *it*-clefts: The myth of exhaustivity. In *Chicago Linguistic Society (CLS) 2014*, vol. 50 [pp. 1, 23, 27, 30, 52, 55, 56, 58, 59, 65, 67, 68, 69, 70, 83, 85, 88, 89, 91, 98, 128, 129, 130, 146, 148, 168, 171].

Prince, Ellen. 1978. A comparison of *wh*-clefts and *it*-cleft in discourse. *Language* 54. 883–906. https://doi.org/10.2307/413238 [pp. 2, 56].

R Core Team. 2017. *R: A Language And Environment For Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/ [pp. 47, 79].

R Core Team. 2019. *R: A Language And Environment For Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/ [pp. 110, 163].

Reeve, Mathew. 2007. Relatives and pronouns in the English cleft construction. In Richard Breheny & Nikolaos Velegrakis (eds.), *UCL working papers in lingustics*, 157–182. University College London [p. 2].

Reeve, Mathew. 2012. *Clefts and their relatives*. John Benjamins. https://doi.org/10.1075/la.185 [pp. 2, 60].

Reinhart, Tanya. 1997. Quantifier-scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20. 335–397 [p. 59].

Renans, Agata. 2016a. *Exhaustivtity. on exclusive particles, clefts, and progressive aspect in Ga (Kwa)*. Potsdam, Germany: Universität Potsdam dissertation. https:

//publishup.uni-potsdam.de/files/8950/renans_diss.pdf [pp. 93, 94, 98, 106, 108, 119, 123, 133].

Renans, Agata. 2016b. Modeling the exhaustivity inference of clefts: Evidence from Ga (Kwa). In Nadine Bade & Polina Berezovskaya (eds.), *Sinn und Bedeutung (SuB)*, vol. 20, 569–588 [pp. 22, 93, 94, 106, 108, 119].

Renans, Agata & Joseph P. De Veaugh-Geiss. 2019. Experimental studies on *it*-clefts and predicate interpretation. *Semantics and Pragmatics* 12(11). Authors listed in reverse alphabetical order, 1–50. https://doi.org/10.3765/sp.12.11 [pp. 3, 5, 15, 22, 23, 146].

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics* 49. 91–136 [pp. 8, 98, 171].

Roberts, Craige. 2011. *only*: A case study in projective meaning. In Barbara H. Partee, Michael Glanzberg & Jurgis Skilters (eds.), *Baltic international yearbook of cognition, logic and communication*, vol. 6, 1–59. Manhattan, KS: New Prairie Press. https://doi.org/10.4148/biyclc.v6i0.1581 [p. 8].

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6). 1–69. https://doi.org/10.3765/sp.5.6 [pp. 28, 59, 171].

Romero, Maribel. 2005. Concealed questions and specificational subjects. *Linguistics and Philosophy* 28(6). 687–737. https://doi.org/10.1007/s10988-005-2654-9 [p. 28].

Romoli, Jacopo. 2012. *Soft but strong. Neg-raising, soft triggers, and exhaustification*. Harvard University dissertation [p. 15].

Romoli, Jacopo & Uli Sauerland. 2017. Presuppositions and accommodation. In Anne Barron, Yueguo Gu & Gerard Steen (eds.), *Routledge handbook of pragmatics*. Routledge [p. 121].

Romoli, Jacopo & Florian Schwarz. 2015. An experimental comparison between presupposition and indirect scalar implicatures. In Florian Schwarz (ed.), *Experimental perspectives on presuppositions*, 215–240. Springer. https://doi.org/10.1007/978-3-319-07980-6 [pp. 18, 19, 37, 38, 85, 87].

van Rooij, Robert & Katrin Schulz. 2005. Only: Meaning and implicatures. In Emar Maier, Corien Bary & Janneke Huitink (eds.), *Sinn und Bedeutung (SuB)*, vol. 9. Nijmegen: Nijmegen Centre for Semantics [p. 8].

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116 [p. 132].

Rooth, Mats. 1996. On the interface principles for intonational focus. In Teresa Galloway & Justin Spence (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 6, 202–226. Ithaca, NY: Cornell University [pp. 27, 56, 89].

Rooth, Mats. 1999. Association with focus or association with presupposition. In Peter Bosch & Rob van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 232–244. Cambridge University Press [p. 27].

Saah, Kofi K. 1994. *Studies in Akan syntax, aquisition and sentence processing*. University of Otawa dissertation [p. 133].

Saah, Kofi K. 2010. Relative clauses in Akan. In Enoch O.. Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 91–107 [p. 133].

van der Sandt, Rob. 1989. Anaphora and accommodation. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and contextual expression*. Dordrecht: Foris [p. 56].

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4). 333–377. https://doi.org/10.1093/jos/9.4.333 [p. 56].

Saur, Eva-Maria. 2013. *Clefts: Discourse function and the nature of exhaustivity violation effects*. Universität Potsdam MSc thesis [pp. 71, 87, 139].

Schad, Daniel J., Michael Betancourt & Shravan Vasishth. 2019. Toward a principled Bayesian workflow in cognitive science. Manuscript; Submitted Arxiv on 29 Apr 2019 (v1), last revised 2 May 2019 (v2). https://arxiv.org/abs/1904.12765v2 [p. 164].

Schad, Daniel J., Shravan Vasishth, Sven Hohenstein & Reinhold Kliegl. 2019. How to capitalize on *a priori* contrasts in linear (mixed) models: A tutorial. Manuscript; Submitted Arxiv on 27 July 2018 (v1), last revised 1 April 2019 (v3). https://arxiv.org/abs/1807.10451v3 [p. 164].

Schlenker, Philippe. 2008. Be articulate: A pragmatic theory of presupposition projection. *Theoretical Linguistics* 34. 157–212 [p. 12].

Schulz, Katrin & Robert Van Rooij. 2006. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* 29(2). 205–250 [p. 129].

Schwarz, Florian. 2009. *Two types of definites in natural language*. Amherst, MA: GLSA [pp. 33, 60, 62].

Schwarz, Florian. 2014. Presuppositions are fast, whether hard or soft – evidence from the visual world. In Todd Snider, Sarah D'Antonio & Mia Weigand (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 24, 1–24. LSA & CLC Publications. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/download/24.1/2139 [pp. 74, 87].

Schwarz, Florian. 2015. Presuppositions vs. asserted content in online processing. In Florian Schwarz (ed.), *Experimental perspectives on presuppositions*, 89–108. Springer. https://doi.org/10.1007/978-3-319-07980-6 [pp. 74, 87].

Schwarzschild, Roger. 1993. Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics* 2(3). 201–248. https://doi.org/10.1007/bf01256743 [p. 96].

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer [p. 18].

Selkirk, Elisabeth O. 1984. *Phonology and syntax. The relation between sound and structure*. Cambridge, MA: MIT Press [p. 133].

Shinners, Pete. 2011. *PyGame [Computer Application]*. http://pygame.org/ [pp. 39, 152].

Šimík, Radek & Christoph Demian. 2019. Definiteness, uniqueness, and maximality in languages with and without articles. Manuscript, Humboldt-Universität zu Berlin. Version from 20 February 2019 cited. [pp. 12, 18, 19].

Simons, Mandy. 2001. On the conversational basis of some presuppositions. In Rachel Hastings, Brendan Jackson & Zsofia Zvolensky (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 11. Ithaca, NY: CLC Publications. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/download/3099/2822 [p. 12].

Simons, Mandy, David Beaver, Judith Tonhauser & Craige Roberts. 2010. What projects and why. In Nan Li & David Lutz (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 20, 309–327. https://doi.org/10.3765/salt.v20i0.2584 [pp. 34, 35].

Singh, Raj, Evelina Fedorenko, Kyle Mahowald & Edward Gibson. 2016. Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science* 40. 607–634. https://doi.org/10.1111/cogs.12260 [p. 17].

Skopeteas, Stavros & Gisbert Fanselow. 2011. Focus and the exclusion of alternatives: On the interaction of syntactic structure with pragmatic inference. *Lingua* 11. 1693–1706. https://doi.org/10.1016/j.lingua.2011.05.005 [pp. 55, 66, 71, 72].

Soames, Scott. 1979. A projection problem for speaker presuppositions. *Linguistic Inquiry* 10. 623–666 [p. 15].

Soames, Scott. 1989. Presupposition. In Dov Gabbay & Franz Guenther (eds.), *Handbook of philosophical logic*, vol. IV, 553–616. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-1171-0_9 [p. 56].

Spector, Benjamin. 2014. Global positive polarity items and obligatory exhaustivity. *Semantics and Pragmatics* 7(11). 1–61. https://doi.org/10.3765/sp.7.11 [p. 72].

Spector, Benjamin & Yasutada Sudo. 2017. Presupposed ignorance and exhaustification: How scalar implicatures and presuppositions interact. *Linguistics and Philosophy* 40(5). 473–517. https://doi.org/10.1007/s10988-017-9208-9 [p. 124].

Spenader, Jennifer. 2002. *Presuppositions in spoken discourse*. Stockholm University dissertation [p. 13].

Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–214. Reprinted in Stalnaker 1999, pages 47–62. New York University Press [pp. 8, 13].

Stan Development Team. 2018. *RStan: The R Interface To Stan*. R package version 2.17.3. https://mc-stan.org/ [pp. 110, 163].

Storto, Gianluca & Michael K. Tanenhaus. 2005. Are scalar implicatures computed online? In Emar Maier, Corien Bary & Janneke Huitink (eds.), *Sinn und Bedeutung (SuB)*, vol. 9, 431–445. Nijmegen: Nijmegen Centre for Semantics [p. 74].

Strawson, Peter. 1950. On referring. *Mind* 59(235). 320–44. https://doi.org/10.1093/mind/LIX.235.320 [p. 15].

Strawson, Peter. 1954. A reply to Mr. Sellars. *Philosophical Review* 63. 216–23 [pp. 16, 17].

Strawson, Peter. 1964. Identifying reference and truth values. *Theoria* 30(2). 96–118. https://doi.org/10.1111/j.1755-2567.1964.tb00404.x [p. 15].

Syrett, Kristen & Julien Musolino. 2013. Collectivity, distributivity, and the interpretation of numerical expressions in child and adult language. *Language Acquisition* 20(4). 259–291. https://doi.org/10.1080/10489223.2013.828060 [p. 117].

Szabo, Zoltan. 2000. Descriptions and uniqueness. *Philosophical Studies* 101. 29–57 [pp. 12, 28].

Szabo, Zoltan. 2004. Descriptions without uniqueness. *Philosophical Studies* [p. 12].

Szabolcsi, Anna. 1994. All quantifiers are not equal: The case of focus. *5th Symposium on Logic and Language. Special issue of Acta Linguistica Hungarica 42–43*. 171–187 [pp. 27, 32, 57].

Szabolcsi, Anna & Bill Haddican. 2004. Conjunction meets negation: A study in cross-linguistic variation. *Journal of Semantics* 21(3). 219–249. https://doi.org/10.1093/jos/21.3.219 [p. 103].

Tiemann, Sonja. 2014. *The processing of* wieder *('again') and other presupposition triggers*. Eberhard Karls Universität Tübingen dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/57169 [pp. 15, 16, 19, 57].

Tiemann, Sonja, Mareike Schmid, Nadine Bade, Bettina Rolke, Ingo Hertrich, Hermann Ackermann, Julia Knapp & Sigrid Beck. 2011. Pyscholinguistic evidence for presuppositions: On-line and off-line data. In Ingo Reich, Eva Horch & Dennis Pauly (eds.), *Sinn und Bedeutung (SuB)*, vol. 15, 581–595. Saarbrücken, Germany: Saarland University Press. http://universaar.uni-saarland.de/monographien/volltexte/2011/30/artikel/Tiemann_et_al_sub15.pdf [pp. 15, 74].

Tieu, Lyn & Manuel Križ. 2017. Connecting the exhaustivity of clefts and the homogeneity of plural definite descriptions in acquisition. In Maria LaMendola & Jennifer Scott (eds.), *41st Boston University Conference on Language Development*, 651–664 [p. 72].

Titov, Elena. 2019. Morphosyntactic encoding of information structure in Akan. *Glossa: A Journal of General Linguistics* 4(1). 27, 1–36. https://doi.org/10.5334/

gjgl.576 [pp. 2, 7, 24, 127, 128, 129, 132, 133, 134, 137, 138, 139, 141, 145, 146, 160].

Tomlinson, John, Todd M. Bailey & Lewis Bott. 2013. Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language* (69). 18–35 [p. 73].

Tonhauser, Judith, David Beaver, Craige Roberts & Mandy Simons. 2013. Towards a taxonomy of projective content. *Language* 89(1). 66–109. https://doi.org/10.1353/lan.2013.0001 [pp. 34, 35, 54, 85].

Vasishth, Shravan, Bruno Nicenboim, Mary E. Beckman, Fangfang Li & Eun Jong Kong. 2018. Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics* 71. 147–161. https://doi.org/10.1016/j.wocn.2018.07.008 [pp. 110, 163, 164].

Velleman, Dan Bridges, David Beaver, Emilie Destruel, Dylan Bumford, Edgar Onea & Liz Coppock. 2012. *It*-clefts are IT (Inquiry Terminating) constructions. In Anca Chereches (ed.), *Semantics and Linguistic Theory (SALT)*, vol. 22, 441–460. https://doi.org/10.3765/salt.v22i0.2640 [pp. 4, 5, 9, 13, 22, 28, 34, 66, 69, 93, 94, 95, 98, 99, 103, 104, 108, 121, 125, 129, 142, 147].

Welby, Pauline. 2006. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34. 343–371. https://doi.org/10.1016/j.wocn.2005.09.001 [p. 77].

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20. 399–467 [p. 59].

Xue, Jingyang & Edgar Onea. 2011. Correlation between presupposition projection and at-issueness: An empirical study. In *ESSLLI 2011 Workshop on Projective Meaning*. Ljubljana, Slovenia [pp. 35, 71, 129].

Zerbian, Sabine. 2007. The subject/object-asymmetry in Northern Sotho. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 323–345. Amsterdam: John Benjamins [p. 138].

Zimmermann, Malte, Joseph P. De Veaugh-Geiss, Swantje Tönnis & Edgar Onea. to appear. (Non-)exhaustivity in focus partitioning: A cross-linguistic investigation on exhaustivity. In: Proceedings of ICHS 2017 [pp. 71, 134, 146, 151, 171].

List of Figures

List of Tables

Appendix A

Chapter 1: "That's not quite it: An experimental investigation of (non-)exhaustivity in clefts"

A.1 Supplementary materials

Data and materials can be found online at https://static.semprag.org/sp.11.3s.zip.

Appendix B

Chapter 2: "On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts"

B.1 Sample stimuli for English experimental task

1. A red-head boy is eating an ice cream.
2. A red-head boy is chasing a butterfly.
3. A red-head child is pushing a toy-car.
4. A red-head child is kicking a ball.
5. A blond girl is spinning a pinwheel.
6. A blond girl is smelling a flower.
7. An old teacher is throwing a book.
8. An old teacher is cutting an apple.
9. A young fireman is moving a hose.
10. A young fireman is climbing a ladder.
11. A tall man is taking a picture.
12. A tall man is reading a newspaper.
13. A small woman is watching a movie.
14. A small woman is brushing her hair.

15. A young policeman is writing a ticket.
16. A young policeman is eating a doughnut.
17. A happy cowboy is riding a horse.
18. A happy cowboy is smoking a cigar.
19. An old sailor is catching a fish.
20. An old sailor is playing the flute.

B.2 Sample stimuli for French experimental task

1. Un enfant roux mange une glace.
2. Un enfant roux poursuit un papillon.
3. Un enfant roux pousse une petite voiture.
4. Un enfant roux tape dans un ballon.
5. Une fille blonde fait tourner un moulinet.
6. Une fille blonde sent une fleur.
7. Une vieille professeur jette un livre.
8. Une vieille professeur coupe une pomme.
9. Un jeune pompier bouge un tuyau.
10. Un jeune pompier monte à une échelle.
11. Un homme grand prend une photo.
12. Un homme grand lit un journal.
13. Une petite femme regarde un film.
14. Une petite femme se brosse les cheveux.
15. Un jeune policier écrit une contravention.
16. Un jeune policier mange un beignet.
17. Un cowboy heureux fait du cheval.
18. Un cowboy heureux fume un cigare.
19. Un vieux marin pêche un poisson.
20. Un vieux marin joue de la flûte.

Appendix C

Chapter 3: "Experimental studies on *it*-clefts and predicate interpretation"

C.1 Instructions to the participants

> Please note that this is not a grammar test! We are not interested in judgments based on such things as "I learned in school that this is correct English, and therefore this sentence must be acceptable." For us it is very important that every answer is based on your own intuition about the acceptability of the sentences in context. There are no right or wrong responses. Keep in mind that it does not help us if you ask someone else for his or her judgment. We are only interested in your opinion.
>
> Even if many sentences may appear to be similar, it is very important that you judge each sentence on its own without letting prior responses influence you. It is possible that some sentences which do not sound acceptable can be improved with a small change. Please do not "correct" the sentences. Each sentence was written that way in order to investigate a specific aspect of English. It is enough that you express your opinion about the acceptability of the sentences as they appear in context without any modifications made to them.

C.2 Target items: Experiment 1 and Experiment 2

In this section we provide a full list of the written stimuli as in Experiment 1, which is broken down into two subsections: Appendix C.2.1 presents the 20 target items with an episodic interpretation, and Appendix C.2.2 presents the 20 target items with

a habitual interpretation. Note that only target items with an episodic interpretation were used in Experiment 2 and Experiment 3.

C.2.1 Episodic (Experiments 1 and 2)

As discussed for Experiment 2 in Section 3.3.2.1 under "Materials", for the auditory version of the experiments the order of conjuncts in the second clause for the factor *Sentence* was reversed for the odd-numbered episodic target items listed below (i.e., 1, 3, 5, 7, 9, 11, 13, 15, 17, 19). For example, for the second sentence in item 01 the written stimuli had the order *She gave birth to* ***Jacob and Ryan***, shown below, whereas the auditory stimuli had the reverse order *She gave birth to* ***Ryan and Jacob***, not shown here. This was done in order to keep the prosodic contours consistent across the items. In every other respect the stimuli were identical. For the odd-numbered episodic target items, by contrast, the order of conjuncts was unchanged across the experiments. Since reconstructing the order of conjuncts for the auditory stimuli is straightforward, we only present the written version of the items here.

01 CONTEXT

Distributive: Jacob and Ryan are Maria's children. Jacob is fifteen and Ryan is two.
Non-Distributive: Jacob and Ryan are twins, sons of Maria.
SENTENCE
it-*Cleft*: It's not Ryan Maria gave birth to. She gave birth to Jacob and Ryan.
SVO: Maria didn't give birth to Ryan. She gave birth to Jacob and Ryan.

02 CONTEXT

Distributive: Marc died in a car accident last year, and Anthony, a fireman, died last week while rescuing people.
Non-Distributive: Marc and Anthony were firemen. They died last week while rescuing people.
SENTENCE
it-*Cleft*: It wasn't Marc who died. Marc and Anthony died.
SVO: Marc didn't die. Marc and Anthony died.

03 CONTEXT

Distributive: Sophia and Kathrine share a car. Sophia drove the car on Monday and Kathrine drove the car on Friday.
Non-Distributive: Sophia and Kathrine tested a new invention of the automotive industry: a car that is driven by two people simultaneously.
SENTENCE
it-*Cleft*: It wasn't Sophia who drove the car. Kathrine and Sophia drove the car.
SVO: Sophia didn't drive the car. Kathrine and Sophia drove the car.

04 CONTEXT

Distributive: Emma and Samuel are causing a lot of trouble! Last week, Emma swallowed a coin on Monday and Samuel swallowed a coin on Wednesday.
Non-Distributive: Emma and Samuel are causing a lot of trouble! Last week, they miraculously managed to break a coin and swallow it (each one half).
SENTENCE
it-*Cleft*: It wasn't Emma who swallowed a coin. Emma and Samuel swallowed a coin.
SVO: Emma didn't swallow a coin. Emma and Samuel swallowed a coin.

05 CONTEXT
Distributive: Isabella is a monogamist. She married her first husband, James, in 1985 and her second husband, Tyler, in 2001.
Non-Distributive: Isabella is a polygamist. She married her husbands, James and Tyler, in a small wedding ceremony in July 2015.
SENTENCE
it-*Cleft*: It wasn't James Isabella married. She married Tyler and James.
SVO: Isabella didn't marry James. She married Tyler and James.

06 CONTEXT
Distributive: Madison and Abigail are marathon runners. Madison won the New York marathon in 2001 and Abigail won the Boston marathon in 2005.
Non-Distributive: Madison and Abigail are marathon runners. Last year, they both ran the New York marathon and they passed the finish line at exactly the same time.
SENTENCE
it-*Cleft*: It wasn't Madison who won a marathon. Madison and Abigail won a marathon.
SVO: Madison didn't win a marathon. Madison and Abigail won a marathon.

07 CONTEXT
Distributive: Chloe is an astronomer. In 2000, she was writing a paper on the chemical reactions on the sun's surface, so she observed the sun. Recently, she was interested in the geology of the moon, so she observed the moon.
Non-Distributive: Chloe is an astronomer. She tried to answer the question about what happens in the atmosphere when both the sun and the moon are visible at the same time. For this reason she made several observations of this phenomenon.
SENTENCE
it-*Cleft*: It wasn't the sun Chloe observed. She observed the moon and the sun.
SVO: Chloe didn't observe the sun. She observed the moon and the sun.

08 CONTEXT
Distributive: Owen is not a lucky man. He proposed to Alice and he was rejected. So he proposed to Linda and he was rejected as well.
Non-Distributive: Owen thought he would be the happiest man in the world if he could marry both Alice and Linda. He wanted to give it a try and one winter evening he popped the question to both of them.
SENTENCE
it-*Cleft*: It wasn't Alice Owen proposed to. He proposed to Alice and Linda.
SVO: Owen didn't propose to Alice. He proposed to Alice and Linda.

09 CONTEXT

Distributive: Kathy and Daniel are judges. In 1995 Kathy sentenced Bob to two years in prison and in 2003 Daniel sentenced Bob to three years in prison.
Non-Distributive: Kathy and Daniel are judges. Last autumn, they were judges on the case of Bob, a drug-dealer. They sentenced him to 10 years in prison.
SENTENCE
it-*Cleft*: It wasn't Kathy who sentenced Bob to prison. Daniel and Kathy sentenced Bob to prison.
SVO: Kathy didn't sentence Bob to prison. Daniel and Kathy sentenced Bob to prison.

10 CONTEXT
Distributive: Carlos and Andrea like biking in a nearby forest. However, they have never seen each other there! Last week, Carlos biked on Monday and Andrea biked on Wednesday.
Non-Distributive: Carlos and Andrea like biking. They own a tandem bike and they use it all the time! Last week, they biked in a nearby forest together.
SENTENCE
it-*Cleft*: It wasn't Carlos who biked. Carlos and Andrea biked.
SVO: Carlos didn't bike. Carlos and Andrea biked.

11 CONTEXT
Distributive: Carter and Mason are a couple. Ava invited Carter to her birthday party and Mason to her graduation party.
Non-Distributive: Carter and Mason are a couple. Ava invited them to her birthday party.
SENTENCE
it-*Cleft*: It wasn't Carter she invited. She invited Mason and Carter.
SVO: She didn't invite Carter. She invited Mason and Carter.

12 CONTEXT
Distributive: Mia went to a party on Monday and she drank gin and tonic. Harper went to a party on Tuesday and she also drank gin and tonic.
Non-Distributive: Mia and Harper went to a party last night. They did not have too much money so they just shared one gin and tonic.
SENTENCE
it-*Cleft*: It wasn't Mia who drank gin and tonic. Mia and Harper drank gin and tonic.
SVO: Mia didn't drink gin and tonic. Mia and Harper drank gin and tonic.

13 CONTEXT
Distributive: Zoe and Emily are sisters. Zoe visited their grandmother on Wednesday and Emily visited her on Friday.
Non-Distributive: Zoe and Emily are sisters. On Wednesday, they went to visit their grandmother together.
SENTENCE
it-*Cleft*: It wasn't Zoe who visited their grandmother. Emily and Zoe visited her.
SVO: Zoe didn't visit their grandmother. Emily and Zoe visited her.

14 CONTEXT
Distributive: Lily and Ethan live together. Last week, Ethan cooked dinner on Tuesday and Lily cooked dinner on Friday.

Non-Distributive: Lily and Ethan live together. Last Friday, Ethan and Lily cooked dinner together.

SENTENCE

it-*Cleft*: It wasn't Ethan who cooked dinner. Ethan and Lily cooked dinner.

SVO: Ethan didn't cook dinner. Ethan and Lily cooked dinner.

15 CONTEXT

Distributive: There was a song competition in Sarah's school. Sarah sang a song on Monday and William sang a song on Wednesday.

Non-Distributive: There was a song competition in Sarah's school. Sarah and her friend William sang a song together.

SENTENCE

it-*Cleft*: It wasn't Sarah who sang a song. William and Sarah sang a song.

SVO: Sarah didn't sing a song. William and Sarah sang a song.

16 CONTEXT

Distributive: Charlotte and Ella are arranging furniture in their new apartment. On Monday Charlotte moved a piano from the bedroom to the living room. On Friday, Ella moved the piano back to the bedroom.

Non-Distributive: Charlotte and Ella are arranging furniture in their new apartment. They are really happy now because when they combined forces they managed to move their heavy piano from the bedroom to the living room.

SENTENCE

it-*Cleft*: It wasn't Charlotte who moved the piano. Charlotte and Ella moved the piano.

SVO: Charlotte didn't move the piano. Charlotte and Ella moved the piano.

17 CONTEXT

Distributive: Olivia and Victoria are colleagues, but they never co-authored a paper. Olivia submitted a new paper in June and Victoria submitted a new paper in July.

Non-Distributive: Olivia and Victoria are colleagues. Recently, they co-authored a paper.

SENTENCE

it-*Cleft*: It wasn't Olivia who wrote a paper. Victoria and Olivia wrote a paper.

SVO: Olivia didn't write a paper. Victoria and Olivia wrote a paper.

18 CONTEXT

Distributive: Noah and Henry love sailing but they never do it together. Last week, Noah sailed on Saturday and Henry sailed on Sunday.

Non-Distributive: Noah and Henry love sailing together. Last Saturday they sailed together again.

SENTENCE

it-*Cleft*: It wasn't Noah who sailed. Noah and Henry sailed.

SVO: Noah didn't sail. Noah and Henry sailed.

19 CONTEXT

Distributive: Anne and Benjamin are wonderful speakers. Anne gave a great speech at Benjamin's wedding in May and Benjamin gave a great speech at Anne's wedding in August.

Non-Distributive: Anne and Benjamin are wonderful speakers. They gave an amazing speech at their best friend's wedding together.

SENTENCE

it-*Cleft*: It wasn't Anne who gave a speech. Benjamin and Anne gave a speech.

SVO: Anne didn't give a speech. Benjamin and Anne gave a speech.

20 CONTEXT

Distributive: Kevin and Susanne feel really adult now. Two months ago Kevin bought an apartment and last month Susanne bought an apartment.

Non-Distributive: Kevin and Susanne feel really adult now. Two months ago they bought an apartment together.

SENTENCE

it-*Cleft*: It wasn't Kevin who bought an apartment. Kevin and Susanne bought an apartment.

SVO: Kevin didn't buy an apartment. Kevin and Susanne bought an apartment.

C.2.2 Habitual (Experiment 1 only)

01 CONTEXT

Distributive: Matthew and Nicholas love swimming. Matthew swims on Mondays and Nicholas swims on Tuesdays.

Non-Distributive: Matthew and Nicholas love swimming. Every Monday they go swimming together.

SENTENCE

it-*Cleft*: It's not Matthew who swims. Matthew and Nicholas swim.

SVO: Matthew doesn't swim. Matthew and Nicholas swim.

02 CONTEXT

Distributive: Patricia plays computer games on the weekends and Martha plays computer games during the week.

Non-Distributive: Every Saturday Patricia and Martha play computer games together.

SENTENCE

it-*Cleft*: It's not Patricia who plays computer games. Martha and Patricia play computer games.

SVO: Patricia does not play computer games. Martha and Patricia play computer games.

03 CONTEXT

Distributive: Dorothy and Richard live together. Dorothy prepares breakfast from Monday to Wednesday and Richard prepares breakfast from Thursday to Sunday.

Non-Distributive: Dorothy and Richard are such a sweet couple! They always prepare their breakfast together.

SENTENCE

it-*Cleft*: It's not Dorothy who prepares breakfast. Dorothy and Richard prepare breakfast.

SVO: Dorothy doesn't prepare breakfast. Dorothy and Richard prepare breakfast.

04 CONTEXT

Distributive: Steve and Carl are good friends but they never go shopping together. Steve loves shopping on Saturdays, but Carl only goes shopping on Mondays.

Non-Distributive: Steve and Carl are good friends and they always go shopping together.
SENTENCE
it-*Cleft*: It's not Steve who goes shopping. Carl and Steve go shopping.
SVO: Steve doesn't go shopping. Carl and Steve go shopping.

05 CONTEXT

Distributive: Scott and Betty are swing dancers, but they never dance together! Whereas Scott dances on Mondays and Tuesday, Betty dances only on Sundays.
Non-Distributive: Scott and Betty are swing dancers and they always dance together.
SENTENCE
it-*Cleft*: It's not Betty who dances. Betty and Scott dance.
SVO: Betty doesn't dance. Betty and Scott dance.

06 CONTEXT

Distributive: Cynthia and Nancy are birdwatchers. Cynthia always watches birds in July and Nina in October.
Non-Distributive: Cynthia and Nancy are birdwatchers and they always do it together.
SENTENCE
it-*Cleft*: It's not Cynthia who watches birds. Nina and Cynthia watch birds.
SVO: Cynthia doesn't watch birds. Nina and Cynthia watch birds.

07 CONTEXT

Distributive: Helen and Kimberly live together. Whenever Helen cooks, Kimberly does the dishes and whenever Kimberly cooks, Helen does the dishes.
Non-Distributive: Helen and Kimberly live together. They have one rule: it doesn't matter who cooks, they always do the dishes together.
SENTENCE
it-*Cleft*: It's not Kimberly who does the dishes. Kimberly and Helen do the dishes.
SVO: Kimberly doesn't do the dishes. Kimberly and Helen do the dishes.

08 CONTEXT

Distributive: Laura and Brian are preparing for their final exams. Whereas Laura always studies in the mornings and keeps her afternoons free, Brian sleeps late and studies in the evenings.
Non-Distributive: Laura and Brian are preparing for their final exams. People think they are crazy because they always study together.
SENTENCE
it-*Cleft*: It's not Laura who studies. Brian and Laura study.
SVO: Laura doesn't study. Brian and Laura study.

09 CONTEXT

Distributive: Larry and Shirley like basketball. Larry plays basketball with his friends just after school and Shirley plays with her sisters on Sundays.
Non-Distributive: Larry and Shirley like basketball. They always play it together.
SENTENCE
it-*Cleft*: It's not Larry who plays basketball. Larry and Shirley play basketball.
SVO: Larry doesn't play basketball. Larry and Shirley play basketball.

10 CONTEXT

Distributive: Once a year Nathalie and Chris go surfing on holidays. Nathalie surfs in Australia in March and Chris surfs in Hawaii in June.

Non-Distributive: Once a year Nathalie and Chris go surfing on holidays together.

SENTENCE

it-*Cleft*: It's not Nathalie who surfs. Chris and Nathalie surf.

SVO: Nathalie doesn't surf. Chris and Nathalie surf.

11 CONTEXT

Distributive: Mike and Colin are travelers. However, they have never traveled together.

Non-Distributive: Mike and Colin are travelers who always travel together.

SENTENCE

it-*Cleft*: It's not Mike who travels. Mike and Colin travel.

SVO: Mike doesn't travel. Mike and Colin travel.

12 CONTEXT

Distributive: Doris and Eric run IT companies. Doris sells computers in California and Eric sells computers in Washington D.C.

Non-Distributive: Doris and Eric run an IT company together which sells computers in California.

SENTENCE

it-*Cleft*: It's not Doris who sells computers. Eric and Doris sell computers.

SVO: Doris doesn't sell computers. Eric and Doris sell computers.

13 CONTEXT

Distributive: Peter and Sandra are film producers. Whereas Peter produces documentaries, Sandra produces action films, so they have never produced anything together.

Non-Distributive: Peter and Sandra produce films together.

SENTENCE

it-*Cleft*: It's not Peter who produces films. Peter and Sandra produce films.

SVO: Peter doesn't produce films. Peter and Sandra produce films.

14 CONTEXT

Distributive: Lisa and Willie are mechanics. They don't like each other so they always work alone.

Non-Distributive: Lisa and Willie are mechanics. They are a great team: together they can repair anything! They never work alone.

SENTENCE

it-*Cleft*: It's not Lisa who repairs washing machines. Willie and Lisa repair washing machines.

SVO: Lisa doesn't repair washing machines. Willie and Lisa repair washing machines.

15 CONTEXT

Distributive: Lawrence and Bill are rivals. They are always quarreling over who bakes the best cupcakes in town.

Non-Distributive: Lawrence and Bill are wonderful. Together they bake the best cupcakes in town.

SENTENCE

it-*Cleft*: It's not Lawrence who bakes cupcakes. Lawrence and Bill bake cupcakes.

SVO: Lawrence doesn't bake cupcakes. Lawrence and Bill bake cupcakes.

16 CONTEXT

Distributive: Nina and Jane are biologists. Nina conducts experiments on frogs and Jane conducts experiments on birds, so they have never cooperated.

Non-Distributive: Nina and Jane are biologists who conduct experiments on frogs together.

SENTENCE

it-*Cleft*: It's not Nina who conducts experiments. Jane and Nina conduct experiments.

SVO: Nina doesn't conduct experiments. Jane and Nina conduct experiments.

17 CONTEXT

Distributive: Carol and Jeffrey are working parents. Carol brings the children to kindergarten from Monday to Wednesday and Jeffrey from Thursday to Friday.

Non-Distributive: Carol and Jeffrey are working parents, but nevertheless they always bring their children to kindergarten together.

SENTENCE

it-*Cleft*: It's not Carol who brings the children to kindergarten. Carol and Jeffrey bring the children to kindergarten.

SVO: Carol doesn't bring the children to kindergarten. Carol and Jeffrey bring the children to kindergarten.

18 CONTEXT

Distributive: Louis and Martin have a beautiful garden. Louis works in the garden on Tuesdays and Thursdays and Martin on Fridays and Saturdays.

Non-Distributive: Louis and Martin have a beautiful garden. On the weekends they always work in their garden together.

SENTENCE

it-*Cleft*: It's not Louis who works in the garden. Martin and Louis work in the garden.

SVO: Louis doesn't work in the garden. Martin and Louis work in the garden.

19 CONTEXT

Distributive: Albert and Gloria are burglars. Albert robs banks in New York and Gloria robs banks in Los Angeles.

Non-Distributive: Albert and Gloria are burglars. They always work by robbing banks together.

SENTENCE

it-*Cleft*: It's not Albert who robs banks. Albert and Gloria rob banks.

SVO: Albert doesn't rob banks. Albert and Gloria rob banks.

20 CONTEXT

Distributive: Ralph and Janice love taking baths. They bath regularly but they never do it together!

Non-Distributive: Ralph and Janice love taking baths. They bath regularly and they always do it together!

SENTENCE

it-*Cleft*: It's not Ralph who takes baths. Janice and Ralph take baths.

SVO: Ralph doesn't take baths. Janice and Ralph take baths.

C.3 Sample filler items

C.3.1 *wh*-Clefts

F01 [*wh*-cleft, +NEG, acceptable]

a. *Context:* Diana is spending her holidays in California and Gary is spending his holidays in Texas.
b. Where Diana and Gary are spending their holidays is not Canada. They're spending their holidays in the USA.

F02 [*wh*-cleft, +NEG, unacceptable]

a. *Context:* Tracy is interested in physics and Dale is interested in medieval literature.
b. What Tracy is interested in isn't medieval literature. She is interested in chemistry.

F03 [*wh*-cleft, –NEG, acceptable]

a. *Context:* Hazel and Randall are architects. Randall plans family homes and Hazel is specialized in skyscrapers.
b. What Randall plans is family homes.

F04 [*wh*-cleft, –NEG, unacceptable]

a. *Context:* Rosa isn't a good skier but she snowboards very well. Tom, on the other hand, is a great skier but is a very poor snowboarder.
b. What Rosa and Tom do best is snowboarding.

C.3.2 Expletive sentences

F05 [expletive, +NEG, acceptable]

a. *Context:* Bradley went shopping on Monday and bought a new pair of sandals. Walter went shopping on Wednesday and also bought a new pair of sandals.
b. It's clear that Bradley and Walter didn't buy new ties. They bought new sandals.

F06 [expletive, +NEG, unacceptable]

a. *Context:* Kenneth and Brenda went to a toy shop with their parents and they could pick out whatever they wanted. Kenneth chose a toy car and Brenda chose a computer game.
b. It's obvious that Brenda didn't choose a doll. She chose a toy car.

F07 [expletive, –NEG, acceptable]

a. *Context:* George gave a radio interview in which he recommended two museums to visit: MoMA in New York and The Louvre in Paris.
b. It's obvious that George recommended MoMA and The Louvre.

F08 [expletive, –NEG, unacceptable]

a. *Context:* Lois is so British: she celebrates five o'clock tee and she never has coffee.
b. It's clear that Lois drinks coffee at 5pm.

C.3.3 *it*-Clefts/Definite pseudoclefts

Context: Michael is on his favorite social network each and every day.

F09 It's a photo that Michael posted and he posted a video.

F10 It's a ransom note that Michael posted and he posted a video.

F11 The thing that Michael posted is a photo and he posted a video.

F12 The thing that Michael posted is a ransom note and he posted a video.

C.4 Target items: Experiment 3

01 see episodic target item 10 (auditory version)

02 see episodic target item 14 (auditory version)

03 see episodic target item 15 (auditory version)

04 see episodic target item 16 (auditory version)

05 see episodic target item 17 (auditory version)

06 see episodic target item 19 (auditory version)

07 see episodic target item 20 (auditory version)

08 CONTEXT

Distributive: Liam and Noah were helping their friend Anna move. They divided the work between them: Liam carried a wardrobe from Anna's old apartment to the truck and Noah carried it from the truck to Anna's new

Non-Distributive: Liam and Noah were helping their friend Anna move last weekend. Anna's wardrobe was so heavy that Liam and Noah had to carry it together. apartment.

SENTENCE

it-*Cleft*: It wasn't Liam who carried the wardrobe. Liam and Noah carried the wardrobe.

SVO: Liam didn't carry the wardrobe. Liam and Noah carried the wardrobe.

09 CONTEXT

Distributive: Ava's car keeps breaking down and she would not be able to start the engine if not for her friends. Last Wednesday, James pushed the car and last Saturday Logan pushed the car.

Non-Distributive: Ava's car broke down last week and it had to be pushed. The car was so heavy that her two friends James and Logan had to push it.

SENTENCE

it-*Cleft*: It wasn't Logan who pushed the car. Logan and James pushed the car.

SVO: Logan didn't push the car. Logan and James pushed the car.

10 CONTEXT

Distributive: Isabella and Sophia are carpenters. Isabella built a beautiful table for her mother and Sophia built a table for her sister.

Non-Distributive: Isabella and Sophia are carpenters. Recently, they built a beautiful table together for their dining room.

SENTENCE

it-*Cleft*: It wasn't Isabella who built a table. Isabella and Sophia built a table.

SVO: Isabella didn't build a table. Isabella and Sophia built a table.

11 CONTEXT

Distributive: Mason and Elijah were betting who is stronger by lifting different objects. Mason lifted the fridge and then Elijah lifted the fridge.

Non-Distributive: Mason and Elijah were betting who is stronger by lifting different objects. None of them managed to lift the fridge alone but they were strong enough to lift it together.

SENTENCE

it-*Cleft*: It wasn't Elijah who lifted the fridge. Elijah and Mason lifted the fridge.

SVO: Elijah didn't lift the fridge. Elijah and Mason lifted the fridge.

12 CONTEXT

Distributive: Helen and Kimberly live together. On Saturday Kimberly did the dishes and yesterday Helen did the dishes.

Non-Distributive: Helen and Kimberly live together. Yesterday Helen cooked, but they did the dishes together.

SENTENCE

it-*Cleft*: It wasn't Kimberly who did the dishes. Kimberly and Helen did the dishes.

SVO: Kimberly didn't do the dishes. Kimberly and Helen did the dishes.

13 CONTEXT

Distributive: Albert and Gloria are burglars. Yesterday, Albert robbed a bank in New York and Gloria robbed a bank in Los Angeles.

Non-Distributive: Albert and Gloria are burglars. Yesterday, they robbed a bank together.

SENTENCE

it-*Cleft*: It wasn't Albert who robbed a bank. Albert and Gloria robbed a bank.

SVO: Albert didn't rob a bank. Albert and Gloria robbed a bank.

14 CONTEXT

Distributive: Jacob and Lucas are brothers and they love baking. Jacob baked a cake for their mother last Monday and Lucas baked a cake for their sister last Saturday.

Non-Distributive: Jacob and Lucas are brothers and they love baking together. Last Saturday they baked a birthday cake for their mother together.

SENTENCE

it-*Cleft*: It wasn't Lucas who baked a cake. Lucas and Jacob baked a cake.

SVO: Lucas didn't bake a cake. Lucas and Jacob baked a cake.

15 CONTEXT

Distributive: Mia and Henry are well organized with their household duties. Mia cleaned the garage last week and Henry cleaned the garage this week.

Non-Distributive: Mia and Henry hate their household duties so they always do them with each other. Last week, Mia and Henry cleaned the garage together.

SENTENCE

it-*Cleft*: It wasn't Mia who cleaned the garage. Mia and Henry cleaned the garage.

SVO: Mia didn't clean the garage. Mia and Henry cleaned the garage.

16 CONTEXT

Distributive: Abigail and Madison are talented math students. Yesterday, they solved on their own and independently from each other the very difficult equation their teacher gave them.

Non-Distributive: Abigail and Madison are talented math students. Yesterday, they joined forces and solved the very difficult equation their teacher gave them together.

SENTENCE

it-*Cleft*: It wasn't Madison who solved the equation. Madison and Abigail solved the equation.

SVO: Madison didn't solve the equation. Madison and Abigail solved the equation.

17 CONTEXT

Distributive: Lisa and Willie are mechanics. Last year Lisa repaired my father's car and this year Willie repaired my father's car.

Non-Distributive: Lisa and Willie are mechanics. Last year they repaired my father's car together.

SENTENCE

it-*Cleft*: It wasn't Lisa who repaired my father's car. Lisa and Willie repaired my father's car.

SVO: Lisa didn't repair my father's car. Lisa and Willie repaired my father's car.

18 CONTEXT

Distributive: Matt and Harper are tailors. Matt sew a summer dress for Alice in 2016 and Harper sew a summer dress for Alice in 2017.

Non-Distributive: Matt and Harper are tailors. Together, they sew a very beautiful summer dress for Alice.

SENTENCE

it-*Cleft*: It wasn't Harper who sew a summer dress for her. Harper and Matt sew a summer dress for her.

SVO: Harper didn't sew a summer dress for her. Harper and Matt sew a summer dress for her.

19 CONTEXT

Distributive: Michael and Amelia love making desserts but they never do it together. Michael made a dessert on Monday and Amelia on Wednesday.

Non-Distributive: Michael and Amelia love making desserts together. Last weekend, they prepared a really delicious dessert for their friend.

SENTENCE

it-*Cleft*: It wasn't Michael who made a dessert. Michael and Amelia made a dessert.

SVO: Michael didn't make a dessert. Michael and Amelia made a dessert.

20 CONTEXT

Distributive: Grace and David are street artists. Last week, Grace painted a mural in San Diego and David painted a mural in Seattle.

Non-Distributive: Grace and David are street artists. Last week, they painted a mural in San Diego together.

SENTENCE

it-*Cleft*: It wasn't David who painted a mural. David and Grace painted a mural.

SVO: David didn't paint a mural. David and Grace painted a mural.

Appendix D

Chapter 4: "*nà*-Cleft (non-) exhaustivity: Variability in Akan"

D.1 Data and analysis

Structure of the data In this section I will provide an overview of the structure of the data. The four factors were as follows:

- Language (*LANG*: AK, DE): between-participant, between-item
- Experiment (*EXP*: EXP1, EXP2): between-participant, (partially) within-item
- Sentence Type (*ST*: DP, CL): within-participant, within-item
- Argument (Akan only) (*ARG*: SBJ, OBJ): within-participant, between-item

Note that since the late manipulations for the Akan version of the experiments were fixed—and thus not all items appeared in the critical combinations [+VER, –EXH] or [+FAL, +CAN], i.e., when the canonical meaning holds but exhaustivity is violated (see Section 4.3.1.3 for details)—the factor *EXP* is only partially a within-item manipulation in Akan. Nevertheless, to capture the shared variance of the items which did appear in both experiments, varying slopes for the factor *EXP* was included in the random effects structure for items in the model reported here (see below). A description of the critical combinations per item in Akan is as follows:

- Items 103, 108, 112, 115, 119, 124, 128, 131 appeared in the critical combinations in both experiments (i.e., [+VER, –EXH] for Experiment I and [+FAL, +CAN] for Experiment II).
- Items 104, 107, 111, 116, 120, 123, 127, 132 appeared in the critical combination [+VER, –EXH] in Experiment I.

- Items 101, 106, 110, 113, 117, 122, 126, 129 appeared in the critical combination [+FAL, +CAN] in Experiment II.
- Items 102, 105, 109, 114, 118, 121, 125, 130 did not appear in the critical combinations in either experiment.
  - Note that items 121 and 130 were ultimately removed from the data set due to errors in the stimuli (see Appendix D.2.2 for removed items).

The data for the target and control conditions for both languages is available at the following links:

- German: https://static.semprag.org/sp.11.3s.zip
- Akan: https://gitup.uni-potsdam.de/deveaugh/public-repository-for-the-paper-na-cleft-non-exhaustivity-variability-in-akan

An overview of the relevant columns of the data frames before data preparation is provided in Table D.1. The column *arg* was added to the German data, but otherwise the information for German is presented as it is found at the above link. All numbers within curly braces indicate the relevant range or list of item- and participant-numbers.

	part	*item*	*expLang*	*exp*	*st*	*arg*
German	(*see † below*)	{1–32}	de	exp1	dp	de.sbj
		{1–32}	de	exp1	cl	de.sbj
	{1.exp2 – 32.exp2}	{1–32}	de	exp2	dp	de.sbj
		{1–32}	de	exp2	cl	de.sbj
Akan	{1.exp1.ak – 31.exp1.ak, 1b.exp1.ak – 7b.exp1.ak} (*see † below*)	{103, 104, 111 ,112, 119, 120, 127, 128}	ak	exp1	dp	ak.sbj
		{107, 108, 115, 116, 123, 124, 131, 132}	ak	exp1	dp	ak.obj
		{103, 104, 111 ,112, 119, 120, 127, 128}	ak	exp1	cl	ak.sbj
		{107, 108, 115, 116, 123, 124, 131, 132}	ak	exp1	cl	ak.obj
	{1.exp2.ak – 31.exp2.ak, 1b.exp2.ak – 7b.exp2.ak} (*see † below*)	{101, 103, 110, 112, 117, 119, 126, 128}	ak	exp2	dp	ak.sbj
		{106, 108, 113, 115, 122, 124, 129, 131}	ak	exp2	dp	ak.obj
		{101, 103, 110, 112, 117, 119, 126, 128}	ak	exp2	cl	ak.sbj
		{106, 108, 113, 115, 122, 124, 129, 131}	ak	exp2	cl	ak.obj

Table D.1 *An overview of the structure of the data used in the analysis.*

† A note on the participants:

- The following participants were removed from the final analysis for unexpected behavior in the exclusive control condition (discussed in Footnote 20 for Akan and Footnote 21 for German):
  - German, Experiment I (verifier): participant '12.exp1'
  - Akan, Experiment I (verifier): participants '22.exp1.ak' and '7b.exp1.ak'
  - Akan, Experiment II (falsifier): participants '8.exp2.ak' and '22.exp2.ak'
- For the German version of Experiment I (verifier), the numbering for the participants does not represent the total number of participants, but rather the scheduling of participants in the lab. That is, when participants signed up they were given a number, and if they later cancelled their assigned number did not end up being represented in the data frame. Thus, the list of the thirty-three participants in Experiment I (verifier) in German is as follows:
  - {1.exp1, 2.exp1, 3.exp1, 4.exp1, 6.exp1, 7.exp1, 10.exp1, 11.exp1, 12.exp1, 14.exp1, 15.exp1, 16.exp1, 17.exp1, 19.exp1, 21.exp1, 22.exp1, 24.exp1, 26.exp1, 27.exp1, 28.exp1, 29.exp1, 31.exp1, 32.exp1, 34.exp1, 35.exp1, 36.exp1, 37.exp1, 38.exp1, 39.exp1, 40.exp1, 41.exp1, 42.exp1, 43.exp1}

  A different scheduling system was used for the remaining experiments, and thus the participant numbering in those cases reflects the total number of participants.
- For the Akan version of the experiments, all participant codes with a 'b' following the number had a 5000 ms delay between uncoverings (e.g., '1b.exp1.ak' had a 5000 ms delay vs. '1.exp1.ak' had a 2000 ms delay); see Section 4.3.1.3 for discussion of the 2000 ms vs. 5000 ms delay.

A general note on the data frame:

- The SVO condition is encoded as 'fo' (which stands for *in situ* focus).

Contrast Coding The contrast coding for the numeric covariates is shown in Table D.2; see Section 4.3.2 (under the paragraph "Bayesian analysis") for a written description of the contrast coding used in the experiment.

	Experimental Conditions											
	AK								DE			
	EXP1				EXP2				EXP1		EXP2	
	DP		CL		DP		CL		DP	CL	DP	CL
	AK.SBJ	AK.OBJ	AK.SBJ	AK.OBJ	AK.SBJ	AK.OBJ	AK.SBJ	AK.OBJ	DE.SBJ	DE.SBJ	DE.SBJ	DE.SBJ
ST	0.5	0.5	–0.5	–0.5	0.5	0.5	–0.5	–0.5	0.5	–0.5	0.5	–0.5
EXP	0.5	0.5	0.5	0.5	–0.5	–0.5	–0.5	–0.5	0.5	0.5	–0.5	–0.5
ST.EXP	0.5	0.5	–0.5	–0.5	–0.5	–0.5	0.5	0.5	0.5	–0.5	–0.5	0.5
LANG	–0.5	–0.5	–0.5	–0.5	–0.5	–0.5	–0.5	–0.5	0.5	0.5	0.5	0.5
ST.LANG	–0.5	–0.5	0.5	0.5	–0.5	–0.5	0.5	0.5	0.5	–0.5	0.5	–0.5
EXP.LANG	–0.5	–0.5	–0.5	–0.5	0.5	0.5	0.5	0.5	0.5	0.5	–0.5	–0.5
ST.EXP.LANG	–0.5	–0.5	0.5	0.5	0.5	0.5	–0.5	–0.5	0.5	–0.5	–0.5	0.5
ARG	0.5	–0.5	0.5	–0.5	0.5	–0.5	0.5	–0.5	0	0	0	0
ST.ARG	0.5	–0.5	–0.5	0.5	0.5	–0.5	–0.5	0.5	0	0	0	0
EXP.ARG	0.5	–0.5	0.5	–0.5	–0.5	0.5	–0.5	0.5	0	0	0	0
ST.EXP.ARG	0.5	–0.5	–0.5	0.5	–0.5	0.5	0.5	–0.5	0	0	0	0

Table D.2 *Contrast coding of the independent variables.*

Model The following binomial mixed logistic regression model was fit using the *brms* package in R; see Section 4.3.2 for details.

```
brm(formula = TVJ ~ 1 +
          ST + EXP + ST.EXP + LANG + ST.LANG +
          EXP.LANG + ST.EXP.LANG + ARG +
          ST.ARG + EXP.ARG + ST.EXP.ARG +
          (1 + ST + ARG + ST.ARG | Participant) +
          (1 + ST + EXP + ST.EXP | Item),
  data = df, family = bernoulli(link='logit'),
  prior = c(set_prior('normal(0, 3)', class = 'Intercept'),
            set_prior('normal(0, 3)', class = 'b'),
            set_prior('normal(0, 3)', class = 'sd'),
            set_prior('lkj(2)', class = 'cor')),
  inits = 0, iter = 4000, cores = 4, chains = 4,
  seed = 42, control = list(adapt_delta = 0.95))
```

Results The model estimates in logit and probability scale with 95% credible intervals are shown in Table D.3. The mean and credible interval in probability scale were calculated from the posterior samples extracted from the model using the `posterior_samples()` function from the *brms* package.

Effects	Posterior mean (logit)	Posterior mean (probability)
ST	–0.12 [–0.66, 0.40]	–0.02 [–0.13, 0.09]
EXP	–0.03 [–1.45, 1.39]	–0.03 [–0.32, 0.26]
ST.EXP	–0.24 [–0.81, 0.32]	–0.13 [-0.38, 0.09]
LANG	0.40 [–0.97, 1.82]	0.07 [–0.21, 0.35]
ST.LANG	–0.01 [–0.54, 0.52]	–0.01 [–0.24, 0.21]
EXP.LANG	0.34 [–1.06, 1.77]	0.14 [–0.43, 0.68]
ST.EXP.LANG	0.47 [–0.08, 1.06]	0.35 [–0.08, 0.82]
ARG	0.02 [–0.75, 0.80]	0.004 [–0.15, 0.15]
ST.ARG	–0.29 [–1.05, 0.46]	–0.09 [–0.39, 0.19]
EXP.ARG	–0.56 [–1.35, 0.21]	–0.20 [–0.51, 0.09]
ST.EXP.ARG	0.29 [–0.47, 1.07]	0.23 [–0.35, 0.84]

Table D.3 *Bayesian analysis showing the mean of the posterior distribution in logit and probability scale plus the 95% credible intervals in brackets.*

D.2 Akan stimuli

In order to ensure the correct focus-background information structure in the stimuli used, for the recording of the auditory materials the Akan speaker read the target sentence in response to a *wh*-question triggering narrow information focus on one of the four fictional roommates; see Section 4.3.1.2 for details. All auditory stimuli can be downloaded at the following GitLab repository: https://gitup.uni-potsdam.de/deveaugh/public-repository-for-the-paper-na-cleft-non-exhaustivity-variability-in-akan.

In preparation of the materials for presentation in this appendix, three typographical errors in the stimuli were identified by a native-speaker consultant. These include the following:

- Item 115: *yinaa* should be *gyinaa* ‘stand’ (Exp. I: Box 4; Exp. II: Box 1)
- Item 118: *tii* should be *tee* ‘pluck’ (Exp. I: Box 4; Exp. II: Box 3)
- Item 119: *nwewe* should be *nwene* ‘weave’ (Exp. I: Box 4; Exp. II: Box 1)

In the stimuli below I have left the text as it appeared in the experiments; moreover, I have left the results for these stimuli in the final analysis, as these typos do not appear to have influenced the truth-value judgments.

D.2.1 Target trials

Item 101 Kwaku na ɔnu nkwan mu.
Kwaku FM RPRON.stir.PRS soup inside
'*It is Kwaku who stirs soups.*'

Experiment I (verifier)

Box 1: Anan
Me si ntomago.
1.SG.SBJ wash.PRS rug
'*I wash rugs.*'

Box 2: Kwaku (+VER)
Me nu nkwan mu.
1.SG.SBJ stir.PRS soup inside
'*I stir soups.*'

Box3: Kwame
Me tutu mpa.
1.SG.SBJ dismount.PRS bed
'*I dismount beds.*'

Box 4: Kofi (+EXH)
Me siesie adwareɛ.
1.SG.SBJ renovate.PRS bathroom
'*I clean bathrooms.*'

Experiment II (falsifier)

Box 1: Kwame
Me tutu mpa.
1.SG.SBJ dismount.PRS bed
'*I dismount beds.*'

Box 2: Kofi (+FAL)
Me nu nkwan mu.
1.SG.SBJ stir.PRS soup inside
'*I stir soups.*'

Box 3: Kwaku (+CAN)
Me nu nkwan mu.
1.SG.SBJ stir.PRS soup inside
'*I stir soups.*'

Box 4: Anan
Me si ntomago.
1.SG.SBJ wash.PRS rug
'*I wash rugs.*'

Item 102 Anan na ɔtɔɔ kɛtɛ.
Anan FM RPRON.buy.PST mat
'*It is Anan who bought a mat.*'

Experiment I (verifier)

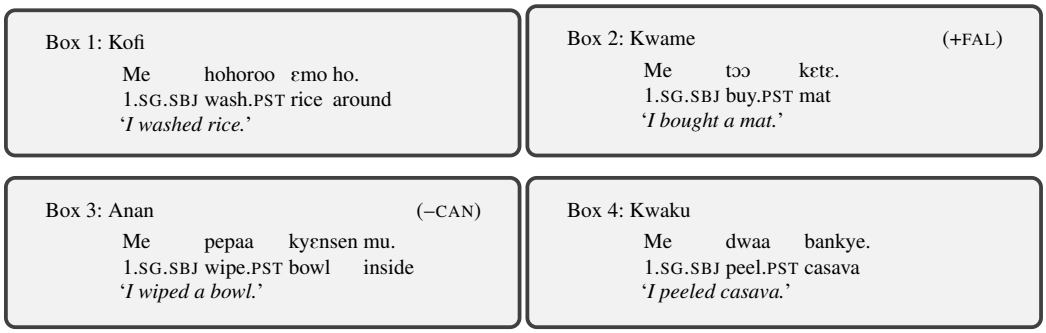
Box 1: Kwaku
Me dwaa bankye.
1.SG.SBJ peel.PST casava
'*I peeled casava.*'

Box 2: Anan (+VER)
Me tɔɔ kɛtɛ.
1.SG.SBJ buy.PST mat
'*I bought a mat.*'

Box 3: Kofi
Me hohoroo ɛmo ho.
1.SG.SBJ wash.PST rice around
'*I washed rice.*'

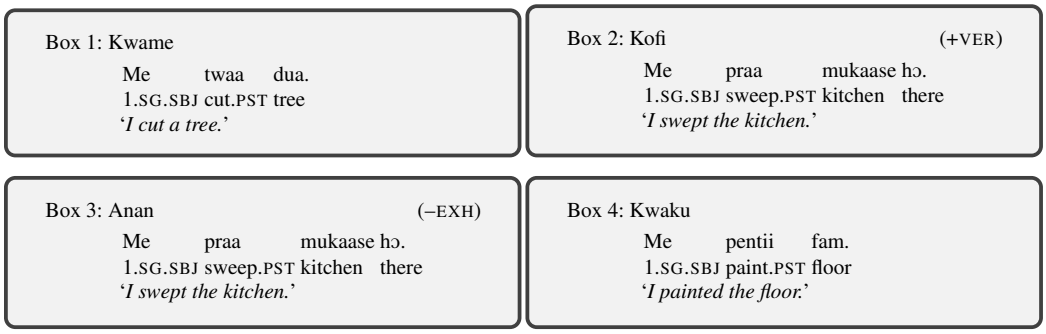
Box 4: Kwame (+EXH)
Me pepaa kyɛnsen mu.
1.SG.SBJ wipe.PST bowl inside
'*I wiped a bowl.*'

Experiment II (falsifier)

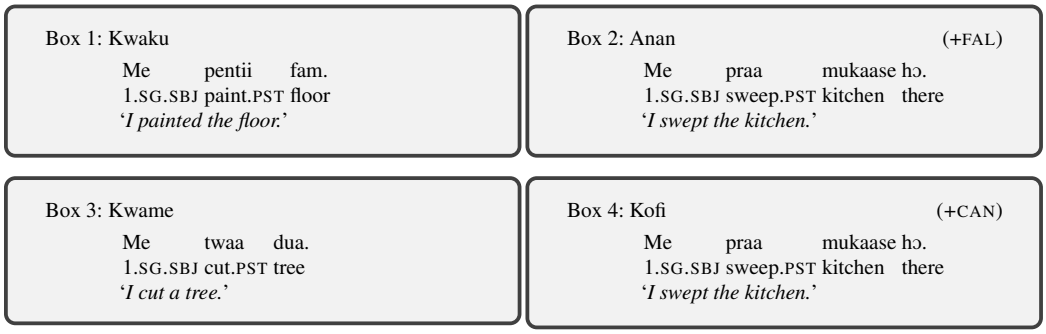


Item 103 Kofi na ɔpraa mukaase hɔ.
Kofi FM RPRON.sweep.PST kitchen there
'*It is Kofi who swept the kitchen.*'

Experiment I (verifier)



Experiment II (falsifier)



Item 104 Kwame na ɔbobɔ ntoma.
Kwame FM RPRON.folds.PRS cloth
'*It is Kwame who folds cloth.*'

Experiment I (verifier)

Box 1: Kofi
Me hata koobi.
1.SG.SBJ dry.PRS tilapia
'*I dry tilapia.*'

Box 2: Kwame (+VER)
Me bobɔ ntoma.
1.SG.SBJ fold.PRS cloth
'*I fold cloth.*'

Box 3: Kwaku
Me pepa mpaboa ho.
1.SG.SBJ polish.PRS shoe around
'*I polish shoes.*'

Box 4: Anan (−EXH)
Me bobɔ ntoma.
1.SG.SBJ fold.PRS cloth
'*I fold cloth.*'

Experiment II (falsifier)

Box 1: Anan
Me pepa mpaboa ho.
1.SG.SBJ polish.PRS shoe around
'*I polish shoes.*'

Box 2: Kwaku (+FAL)
Me bobɔ ntoma.
1.SG.SBJ fold.PRS cloth
'*I fold cloth.*'

Box 3: Kofi
Me hata koobi.
1.SG.SBJ dry.PRS tilapia
'*I dry tilapia.*'

Box 4: Kwame (−CAN)
Me siesie adwareɛ.
1.SG.SBJ renovate.PRS bathroom
'*I clean bathrooms.*'

Item 105 Kofi na Kumi tuu ne fo.
Kofi FM Kumi give.PST 3.SG.OBJ advice
'*It is Kofi to whom Kumi gave advice.*'

Experiment I (verifier)

Box 1: Kwame
Kumi piraa me.
Kumi wound.PST 1.SG.OBJ
'*Kumi wounded me.*'

Box 2: Kofi (+VER)
Kumi tuu me fo.
Kumi give.PST 1.SG.OBJ advice
'*Kumi gave me advice.*'

Box 3: Anan
Kumi gyaa me kwan.
Kumi escort.PST 1.SG.OBJ on.way
'*Kumi escorted me on my way.*'

Box 4: Kwaku (+EXH)
Kumi hyiaa me.
Kumi meet.PST 1.SG.OBJ.
'*Kumi met me.*'

Experiment II (falsifier)

Box 1: Anan
Kumi gyaa me kwan.
Kumi escort.PST 1.SG.OBJ on.way
'*Kumi escorted me on my way.*'

Box 2: Kwaku (+FAL)
Kumi tuu me fo.
Kumi give.PST 1.SG.OBJ advice
'*Kumi gave me advice.*'

Box 3: Kofi (−CAN)
Kumi hyiaa me.
Kumi meet.PST 1.SG.OBJ.
'*Kumi met me.*'

Box 4: Kwame
Kumi piraa me.
Kumi wound.PST 1.SG.OBJ
'*Kumi wounded me.*'

Item 106 Kwame na Akosua kyea no.
Kwame FM Akosua greet.PRS 3.SG.OBJ
'*It is Kwame who Akosua greets.*'

Experiment I (verifier)

Box 1: Kofi
Akosua tia me mu.
Akosua step.on.PRS 1.SG.OBJ in
'*Akosua kicks me.*'

Box 2: Kwame (+VER)
Akosua kyea me.
Akosua greet.PRS 1.SG.OBJ
'*Akosua greets me.*'

Box 3: Kwaku (+EXH)
Akosua kogya me.
Akosua accompany.PRS 1.SG.OBJ
'*Akosua accompanies me.*'

Box 4: Anan
Akosua bo me.
Akosua beat.PRS 1.SG.OBJ
'*Akosua beats me.*'

Experiment II (falsifier)

Box 1: Kwaku
Akosua kogya me.
Akosua accompany.PRS 1.SG.OBJ
'*Akosua accompanies me.*'

Box 2: Anan (+FAL)
Akosua kyea me.
Akosua greet.PRS 1.SG.OBJ
'*Akosua greets me.*'

Box 3: Kwame (+CAN)
Akosua kyea me.
Akosua greet.PRS 1.SG.OBJ
'*Akosua greets me.*'

Box 4: Kofi
Akosua tia me mu.
Akosua step.on.PRS 1.SG.OBJ in
'*Akosua kicks me.*'

Item 107 Kwaku na Adofo sisii no.
Kwaku FM Adofo cheat.PST 3.SG.OBJ
'*It is Kwaku who Adofo cheated.*'

Experiment I (verifier)

Box 1: Anan
Adofo dii me atɛm.
Adofo eat.PST 1.SG.OBJ insult
'*Adofo insulted me.*'

Box 2: Kwaku (+VER)
Adofo sisii me.
Adofo cheat.PST 1.SG.OBJ
'*Adofo cheated me.*'

Box 3: Kwame (−EXH)
Adofo sisii me.
Adofo cheat.PST 1.SG.OBJ
'*Adofo cheated me.*'

Box 4: Kofi
Adofo sɔɔ me so.
Adofo support.PST 1.SG.OBJ top
'*Adofo supported me (financially).*'

Experiment II (falsifier)

Box 1: Kwame
Adofo sɔɔ me so.
Adofo support.PST 1.SG.OBJ top
'*Adofo supported me (financially).*'

Box 2: Kofi (+FAL)
Adofo sisii me.
Adofo cheat.PST 1.SG.OBJ
'*Adofo cheated me.*'

Box 3: Anan
Adofo dii me atɛm.
Adofo eat.PST 1.SG.OBJ insult
'*Adofo insulted me.*'

Box 4: Kwaku (−CAN)
Adofo fumm me mu.
Adofo surprise.PST 1.SG.OBJ inside
'*Adofo surprised me.*'

Item 108 Anan na Panyin gye ne so.
Anan FM Panyin respond.PRS 3.SG.OBJ top
'*It is Anan who Panyin responds to.*'

Experiment I (verifier)

Box 1: Kwaku
Panyin sra me.
Panyin visit.PRS 1.SG.OBJ
'*Panyin visits me.*'

Box 2: Anan (+VER)
Panyin gye me so.
Panyin respond.PRS 1.SG.OBJ top
'*Panyin responds to me.*'

Box 3: Kofi
Panyin po me.
Panyin reject.PRS 1.SG.OBJ
'*Panyin rejects me.*'

Box 4: Kwame (−EXH)
Panyin gye me so.
Panyin respond.PRS 1.SG.OBJ top
'*Panyin responds to me.*'

Experiment II (falsifier)

Box 1: Kofi

Panyin po me.
Panyin reject.PRS 1.SG.OBJ
'*Panyin rejects me.*'

Box 2: Kwame (+FAL)

Panyin gye me so.
Panyin respond.PRS 1.SG.OBJ top
'*Panyin responds to me.*'

Box 3: Kwaku

Panyin sra me.
Panyin visit.PRS 1.SG.OBJ
'*Panyin visits me.*'

Box 4: Anan (+CAN)

Panyin gye me so.
Panyin respond.PRS 1.SG.OBJ top
'*Panyin responds to me.*'

Item 109 Kofi na ɔfra koko.
Kofi FM RPRON.mix.PRS porridge
'*It is Kofi who mixes porridge.*'

Experiment I (verifier)

Box 1: Kwame

Me bu nsuo.
1.SG.SBJ fetch.PRS water
'*I fetch water.*'

Box 2: Kofi (+VER)

Me fra koko.
1.SG.SBJ mix.PRS porridge
'*I mix porridge.*'

Box 3: Anan

Me kata kyɛnsee so.
1.SG.SBJ cover.PRS bowl top
'*I cover bowls.*'

Box 4: Kwaku (+EXH)

Me nom Alomo.
1.SG.SBJ drink.PRS Alomo
'*I drink Alomo* [=local Gin].'

Experiment II (falsifier)

Box 1: Anan

Me kata kyɛnsee so.
1.SG.SBJ cover.PRS bowl top
'*I cover bowls.*'

Box 2: Kwaku (+FAL)

Me fra koko.
1.SG.SBJ mix.PRS porridge
'*I mix porridge.*'

Box 3: Kofi (−CAN)

Me nom Alomo.
1.SG.SBJ drink.PRS Alomo
'*I drink Alomo* [=local Gin].'

Box 4: Kwame

Me bu nsuo.
1.SG.SBJ fetch.PRS water
'*I fetch water.*'

Item 110 Kwame na ɔpepaa akonwa ho.
Kwame FM RPRON.scrub.PST chair around
'*It is Kwame who scrubbed a chair.*'

Experiment I (verifier)

Box 1: Kofi
Me hyee nwura.
1.SG.SBJ burn.PST rubbish.
'*I burned rubbish.*'

Box 2: Kwame (+VER)
Me pepaa akonwa ho.
1.SG.SBJ clean.PST chair around
'*I cleaned a chair.*'

Box 3: Kwaku
Me twaa sini.
1.SG.SBJ film.PST video
'*I filmed a video.*'

Box 4: Anan (+EXH)
Me nyaa abasobɔdeɛ.
1.SG.SBJ win.PST prize
'*I won a prize.*'

Experiment II (falsifier)

Box 1: Kwaku
Me twaa sini.
1.SG.SBJ film.PST video
'*I filmed a video.*'

Box 2: Anan (+FAL)
Me pepaa akonwa ho.
1.SG.SBJ clean.PST chair around
'*I cleaned a chair.*'

Box 3: Kwame (+CAN)
Me pepaa akonwa ho.
1.SG.SBJ clean.PST chair around
'*I cleaned a chair.*'

Box 4: Kofi
Me hyee nwura.
1.SG.SBJ burn.PST rubbish.
'*I burned rubbish.*'

Item 111 Kwaku na ɔtwii kaa.
Kwaku FM RPRON.drive.PST car
'*It is Kwaku who drove a car.*'

Experiment I (verifier)

Box 1: Anan
Me hwɛɛ agyinamoa.
1.SG.SBJ take.care.of.PST cat
'*I took care of a cat.*'

Box 2: Kwaku (+VER)
Me twii kaa.
1.SG.SBJ drive.PST car
'*I drove a car.*'

Box 3: Kwame (–EXH)
Me twii kaa.
1.SG.SBJ drive.PST car
'*I drove a car.*'

Box 4: Kofi
Me kumm aponkye.
1.SG.SBJ kill.PST goat
'*I killed a goat.*'

Experiment II (falsifier)

Box 1: Kwame
Me kumm aponkye.
1.SG.SBJ kill.PST goat
'*I killed a goat.*'

Box 2: Kofi (+FAL)
Me twii kaa.
1.SG.SBJ drive.PST car
'*I drove a car.*'

Box 3: Anan
Me hwɛɛ agyinamoa.
1.SG.SBJ take.care.of.PST cat
'*I took care of a cat.*'

Box 4: Kwaku (–CAN)
Me nyaa abasobɔdeɛ.
1.SG.SBJ win.PST prize
'*I won a prize.*'

Item 112 Anan na ɔkɔ abɛ so.
Anan FM RPRON.go.PRS palm.tree up
'*It is Anan who goes up a palm tree.*'

Experiment I (verifier)

Box 1: Kwaku
Me tu nɔteɛ.
1.SG.SBJ dig.PRS sand
'*I dig in the sand.*'

Box 2: Anan (+VER)
Me kɔ abɛ so.
1.SG.SBJ go.PRS plam.tree up
'*I go up a palm tree.*'

Box 3: Kofi
Me hyɛ ɛkyɛ.
1.SG.SBJ wear.PRS hat
'*I wear a hat.*'

Box 4: Kwame (–EXH)
Me kɔ abɛ so.
1.SG.SBJ go.PRS plam.tree up
'*I go up a palm tree.*'

Experiment II (falsifier)

Box 1: Kofi
Me hyɛ ɛkyɛ.
1.SG.SBJ wear.PRS hat
'*I wear a hat.*'

Box 2: Kwame (+FAL)
Me kɔ abɛ so.
1.SG.SBJ go.PRS plam.tree up
'*I go up a palm tree.*'

Box 3: Kwaku
Me tu nɔteɛ.
1.SG.SBJ dig.PRS sand
'*I dig in the sand.*'

Box 4: Anan (+CAN)
Me kɔ abɛ so.
1.SG.SBJ go.PRS plam.tree up
'*I go up a palm tree.*'

Item 113 Kwaku na Afum da m'ase.
Kwaku FM Afum thank.PRS 3.SG.POSS.under
'*It is Kwaku who Afum thanks.*'

Experiment I (verifier)

Box 1: Anan
Afum tan me.
Afum hate.PRS 1.SG.OBJ
'*Afum hates me.*'

Box 2: Kwaku (+VER)
Afum da m'ase.
Afum thank.PRS 1.SG.POSS.under
'*Afum thanks me.*'

Box 3: Kwame
Afum hu me.
Afum see.PRS 1.SG.OBJ
'*Afum sees me.*'

Box 4: Kofi (+EXH)
Afum huru me.
Afum make.fun.of.PRS 1.SG.OBJ
'*Afum makes fun of me.*'

Experiment II (falsifier)

Box 1: Kwame
Afum hu me.
Afum see.PRS 1.SG.OBJ
'*Afum sees me.*'

Box 2: Kofi (+FAL)
Afum da m'ase.
Afum thank.PRS 1.SG.POSS.under
'*Afum thanks me.*'

Box 3: Kwaku (+CAN)
Afum da m'ase.
Afum thank.PRS 1.SG.POSS.under
'*Afum thanks me.*'

Box 4: Anan
Afum tan me.
Afum hate.PRS 1.SG.OBJ
'*Afum hates me.*'

Item 114 Anan na Ataa tenetenee no.
Anan FM Ataa correct.PST 3.SG.OBJ
'*It is Anan who Ataa corrected.*'

Experiment I (verifier)

Box 1: Kwaku
Ataa yii me ayɛ.
Ataa give.PST 1.SG.OBJ praise
'*Ataa praised me.*'

Box 2: Anan (+VER)
Ataa tenetenee me.
Ataa correct.PST 1.SG.OBJ
'*Ataa corrected me.*'

Box 3: Kofi
Ataa pemm me.
Ataa hit.PST 1.SG.OBJ
'*Ataa hit me.*'

Box 4: Kwame (+EXH)
Ataa gyaa me hɔ.
Ataa leave.PST 1.SG.OBJ there
'*Ataa left me behind.*'

Experiment II (falsifier)

Box 1: Kofi
Ataa pemm me.
Ataa hit.PST 1.SG.OBJ
'*Ataa hit me.*'

Box 2: Kwame (+FAL)
Ataa tenetenee me.
Ataa correct.PST 1.SG.OBJ
'*Ataa corrected me.*'

Box 3: Anan (−CAN)
Ataa gyaa me hɔ.
Ataa leave.PST 1.SG.OBJ there
'*Ataa left me behind.*'

Box 4: Kwaku
Ataa yii me ayɛ.
Ataa give.PST 1.SG.OBJ praise
'*Ataa praised me.*'

Item 115 Kofi na Nsowaa kyerɛɛ n'aseɛ.
Kofi FM Nsowaa teach.PST 3.SG.OBJ'under
'*It is Kofi who Nsowaa explained things to.*'

Experiment I (verifier)

Box 1: Kwame
Nsowaa somaa me.
Nsowaa send.PST 1.SG.OBJ
'*Nsowaa sent me.*'

Box 2: Kofi (+VER)
Nsowaa kyerɛɛ me ase.
Nsowaa explain.PST 1.SG.OBJ under
'*Nsowaa explained things to me.*'

Box 3: Anan (−EXH)
Nsowaa kyerɛɛ me ase.
Nsowaa explain.PST 1.SG.OBJ under
'*Nsowaa explained things to me.*'

Box 4: Kwaku
Nsowaa yinaa me so.
Nsowaa stand.on.PST 1.SG.OBJ on
'*Nsowaa stood on me.*'

Experiment II (falsifier)

Box 1: Anan
Nsowaa yinaa me so.
Nsowaa stand.on.PST 1.SG.OBJ on
'*Nsowaa stood on me.*'

Box 2: Kwaku (+FAL)
Nsowaa kyerɛɛ me ase.
Nsowaa explain.PST 1.SG.OBJ under
'*Nsowaa explained things to me.*'

Box 3: Kwame
Nsowaa somaa me.
Nsowaa send.PST 1.SG.OBJ
'*Nsowaa sent me.*'

Box 4: Kofi (+CAN)
Nsowaa kyerɛɛ me ase.
Nsowaa explain.PST 1.SG.OBJ under
'*Nsowaa explained things to me.*'

Item 116 Kwame na Tawia dɛɛdɛɛ no.
Kwame FM Tawia convince.PRS 3.SG.OBJ
'*It is Kwame who Tawia convinces.*'

Experiment I (verifier)

Box 1: Kofi
Tawia bam me.
Tawia hug.PRS 1.SG.OBJ
'*Tawia hugs me.*'

Box 2: Kwame (+VER)
Tawia dɛɛdɛɛ me.
Tawia convince.PRS 1.SG.OBJ
'*Tawia convinces me.*'

Box 3: Kwaku
Tawia sɔ me mu.
Tawia hold.PRS 1.SG.OBJ in
'*Tawia holds me.*'

Box 4: Anan (–EXH)
Tawia dɛɛdɛɛ me.
Tawia convince.PRS 1.SG.OBJ
'*Tawia convinces me.*'

Experiment II (falsifier)

Box 1: Kwaku
Tawia sɔ me mu.
Tawia hold.PRS 1.SG.OBJ in
'*Tawia holds me.*'

Box 2: Anan (+FAL)
Tawia dɛɛdɛɛ me.
Tawia convince.PRS 1.SG.OBJ
'*Tawia convinces me.*'

Box 3: Kofi
Tawia bam me.
Tawia hug.PRS 1.SG.OBJ
'*Tawia hugs me.*'

Box 4: Kwame (–CAN)
Tawia bo me.
Tawia beat.PRS 1.SG.OBJ
'*Tawia beats me.*'

Item 117 Kwaku na ɔyi prɛte mu.
Kwaku FM RPRON.clean.PRS plate inside
'*It is Kwaku who cleans plates.*'

Experiment I (verifier)

Box 1: Anan
Me kyekyɛ ɛnam.
1.SG.SBJ share.PRS meat
'*I share meat.*'

Box 2: Kwaku (+VER)
Me yi prɛte mu.
1.SG.SBJ clean.PRS plate inside
'*I clean plates.*'

Box 3: Kwame
Me yɛ abomu.
1.SG.SBJ prepare.PRS stew
'*I prepare stew.*'

Box 4: Kofi (+EXH)
Me noa kosua.
1.SG.SBJ boil.PRS egg
'*I boil eggs.*'

Experiment II (falsifier)

Box 1: Kwame Me yɛ abomu. 1.SG.SBJ prepare.PRS stew '*I prepare stew.*'	Box 2: Kofi (+VER) Me yi prɛte mu. 1.SG.SBJ clean.PRS plate inside '*I clean plates.*'
Box 3: Kwaku (+CAN) Me yi prɛte mu. 1.SG.SBJ clean.PRS plate inside '*I clean plates.*'	Box 4: Anan Me kyɛkyɛ ɛnam. 1.SG.SBJ share.PRS meat '*I share meat.*'

Item 118 Anan na ɔtoo ataadeɛ so.
Anan FM RPRON.iron.PST shirt on
'*It is Anan who ironed a shirt.*'

Experiment I (verifier)

Box 1: Kwaku Me dɔɔ ewura. 1.SG.SBJ weed.PST bush '*I weeded a bush.*'	Box 2: Anan (+VER) Me too ataadeɛ so. 1.SG.SBJ iron.PST shirt on '*I ironed a shirt.*'
Box 3: Kofi Me drɔɔ mfonini. 1.SG.SBJ draw.PST picture '*I drew a picture.*'	Box 4: Kwame (+EXH) Me tii nhwiren. 1.SG.SBJ pluck.PST flower '*I plucked a flower.*'

Experiment II (falsifier)

Box 1: Kofi Me dɔɔ ewura. 1.SG.SBJ weed.PST bush '*I weeded a bush.*'	Box 2: Kwame (+FAL) Me too ataadeɛ so. 1.SG.SBJ iron.PST shirt on '*I ironed a shirt.*'
Box 3: Anan (−CAN) Me tii nhwiren. 1.SG.SBJ pluck.PST flower '*I plucked a flower.*'	Box 4: Kwaku Me drɔɔ mfonini. 1.SG.SBJ draw.PST picture '*I drew a picture.*'

Item 119 Kofi na ɔtɔn kasatrofoɔ.
Kofi FM RPRON.sell.PRS mobile.phone
'*It is Kofi who sells mobile phones.*'

Experiment I (verifier)

Box 1: Kwame
Me pam fugu.
I sew.PRS fugu
'*I sew fugu* [=traditional dress].'

Box 2: Kofi (+VER)
Me tɔn kasatrofoɔ.
I sell.PRS mobile.phone
'*I sell mobile phones.*'

Box 3: Anan (−EXH)
Me tɔn kasatrofoɔ.
I sell.PRS mobile.phone
'*I sell mobile phones.*'

Box 4: Kwaku
Me nwewe kente.
I weave.PRS kente
'*I weave kente* [=type of weaving].'

Experiment II (falsifier)

Box 1: Anan
Me nwewe kente.
I weave.PRS kente
'*I weave kente* [=type of weaving].'

Box 2: Kwaku (+FAL)
Me tɔn kasatrofoɔ.
I sell.PRS mobile.phone
'*I sell mobile phones.*'

Box 3: Kwame
Me pam fugu.
I sew.PRS fugu
'*I sew fugu* [=traditional dress].'

Box 4: Kofi (+CAN)
Me tɔn kasatrofoɔ.
I sell.PRS mobile.phone
'*I sell mobile phones.*'

Item 120 Kwame na ɔtoo pono mu.
Kwame FM RPRON.shut.PST gate inside
'*It is Kwame who shut a gate.*'

Experiment I (verifier)

Box 1: Kofi
Me soaa atwedeɛ.
1.SG.SBJ carry.PST ladder
'*I carried a ladder.*'

Box 2: Kwame (+VER)
Me too pono mu.
1.SG.SBJ shut.PST gate inside
'*I shut a gate.*'

Box 3: Kwaku
Me twitwii tiefi so.
1.SG.SBJ clean.PST toilet on
'*I cleaned a toilet.*'

Box 4: Anan (−EXH)
Me too pono mu.
1.SG.SBJ shut.PST gate inside
'*I shut a gate.*'

Experiment II (falsifier)

Box 1: Kwaku
Me twitwii tiefi so.
1.SG.SBJ clean.PST toilet on
'*I cleaned a toilet.*'

Box 2: Anan (+FAL)
Me too pono mu.
1.SG.SBJ shut.PST gate inside
'*I shut a gate.*'

Box 3: Kofi
Me soaa atwedeɛ.
1.SG.SBJ carry.PST ladder
'*I carried a ladder.*'

Box 4: Kwame (−CAN)
Me munii dunsini.
1.SG.SBJ turn.PST log
'*I turned a log.*'

Item 122 Kwame na Amma tee n'ase.
Kwame FM Amma understand.PST 3.SG.OBJ'inside
'*It is Kwame who Amma understood.*'

Experiment I (verifier)

Box 1: Kofi
Amma wiaa me.
Amma steal.from.PST 1.SG.OBJ
'*Amma stole from me.*'

Box 2: Kwame (+VER)
Amma tee m'ase.
Amma understand.PST 1.SG.OBJ'inside
'*Amma understood me.*'

Box 3: Kwaku
Amma gyinaa me.
Amma stop.PST 1.SG.OBJ
'*Amma stopped me.*'

Box 4: Anan (+EXH)
Amma fumm me mu.
Amma surprise.PST 1.SG.OBJ inside
'*Amma surprised me.*'

Experiment I (verifier)

Box 1: Kwaku
Amma gyinaa me.
Amma stop.PST 1.SG.OBJ
'*Amma stopped me.*'

Box 2: Anan (+FAL)
Amma tee m'ase.
Amma understand.PST 1.SG.OBJ'inside
'*Amma understood me.*'

Box 3: Kwame (+CAN)
Amma tee m'ase.
Amma understand.PST 1.SG.OBJ'inside
'*Amma understood me.*'

Box 4: Kofi
Amma wiaa me.
Amma steal.from.PST 1.SG.OBJ
'*Amma stole from me.*'

Item 123 Kwaku na Mansa si ne mu.
Kwaku FM Mansa hit.PRS 3.SG.OBJ inside
'*It is Kwaku who Mansa hits.*'

Experiment I (verifier)

Box 1: Anan
Mansa kae me.
Mansa remind.PRS 1.SG.OBJ
'*Mansa reminds me.*'

Box 2: Kwaku (+VER)
Mansa si me mu.
Mansa hit.PRS 1.SG.OBJ inside
'*Mansa hits me.*'

Box 3: Kwame (−EXH)
Mansa si me mu.
Mansa hit.PRS 1.SG.OBJ inside
'*Mansa hits me.*'

Box 4: Kofi
Mansa tea me.
Mansa scold.PRS 1.SG.OBJ
'*Mansa scolds me.*'

Experiment II (falsifier)

Box 1: Kwame
Mansa tea me.
Mansa scold.PRS 1.SG.OBJ
'*Mansa scolds me.*'

Box 2: Kofi (+FAL)
Mansa si me mu.
Mansa hit.PRS 1.SG.OBJ inside
'*Mansa hits me.*'

Box 3: Anan
Mansa kae me.
Mansa remind.PRS 1.SG.OBJ
'*Mansa reminds me.*'

Box 4: Kwaku (−CAN)
Mansa huru me.
Mansa make.fun.of.PRS 1.SG.OBJ
'*Mansa makes fun of me.*'

Item 124 Anan na Yaa pamoo no.
Anan FM Yaa sack.PST 3.SG.OBJ
'*It is Anan who Yaa kicked out.*'

Experiment I (verifier)

Box 1: Kwaku
Yaa titii me ho.
Yaa scratch.PST 1.SG.OBJ around
'*Yaa scratched me.*'

Box 2: Anan (+VER)
Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Box 3: Kofi
Yaa hyiraa me.
Yaa bless.PST 1.SG.OBJ
'*Yaa blessed me.*'

Box 4: Kwame (−EXH)
Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Experiment II (falsifier)

Box 1: Kofi
Yaa hyiraa me.
Yaa bless.PST 1.SG.OBJ
'*Yaa blessed me.*'

Box 2: Kwame (+FAL)
Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Box 3: Kwaku
Yaa titii me ho.
Yaa scratch.PST 1.SG.OBJ around
'*Yaa scratched me.*'

Box 4: Anan (+CAN)
Yaa pamoo me.
Yaa sack.PST 1.SG.OBJ
'*Yaa kicked me out.*'

Item 125 Kofi na ɔgyinaa amena mu.
Kofi FM RPRON.stand.in.PST pit inside
'*It is Kofi who stood inside a pit.*'

Experiment I (verifier)

Box 1: Kwame
Me tuu mmirika.
1.SG.SBJ run.PST race
'*I ran a race.*'

Box 2: Kofi (+VER)
Me gyinaa amena mu.
1.SG.SBJ stand.in.PST pit inside
'*I stood inside a pit.*'

Box 3: Anan
Me hwɛɛ yikyerɛ.
1.SG.SBJ watch.PST show
'*I watched a theater show.*'

Box 4: Kwaku (+EXH)
Me suaa kasa.
1.SG.SBJ learn.PST language
'*I learned a language.*'

Experiment II (falsifier)

Box 1: Anan
Me hwɛɛ yikyerɛ.
1.SG.SBJ watch.PST show
'*I watched a theater show.'*

Box 2: Kwaku (+FAL)
Me gyinaa amena mu.
1.SG.SBJ stand.in.PST pit inside
'*I stood inside a pit.*'

Box 3: Kofi (−CAN)
Me suaa kasa.
1.SG.SBJ learn.PST language
'*I learned a language.*'

Box 4: Kwame
Me tuu mmirika.
1.SG.SBJ run.PST race
'*I ran a race.*'

Item 126 Kwame na ɔforo bepɔ.
Kwame FM RPRON.climb.PRS mountain
'*It is Kwame who climbs mountains.*'

Experiment I (verifier)

Box 1: Kofi
Me twa nsuo.
I cross.PRS river
'*I cross rivers.*'

Box 2: Kwame (+VER)
Me foro bepɔ.
I climb.PRS mountain
'*I climb mountains.*'

Box 3: Kwaku
Me we akokɔ.
I chew.PRS chicken
'*I eat chicken.*'

Box 4: Anan (+EXH)
Me muni dunsini.
I turn.PRS log
'*I turn logs.*'

Experiment II (falsifier)

Box 1: Kwaku
Me we akokɔ.
I chew.PRS chicken
'*I eat chicken.*'

Box 2: Anan (+FAL)
Me foro bepɔ.
I climb.PRS mountain
'*I climb mountains.*'

Box 3: Kwame (+CAN)
Me foro bepɔ.
I climb.PRS mountain
'*I climb mountains.*'

Box 4: Kofi
Me twa nsuo.
I cross.PRS river
'*I cross rivers.*'

Item 127 Kwaku na ɔsiesie ɛdan mu.
Kwaku FM RPRON.tidy.up.PRS room inside
'*It is Kwaku who tidies up rooms.*'

Experiment I (verifier)

Box 1: Anan
Me hwie nsa.
1.SG.SBJ pour.PRS drink
'*I pour drinks.*'

Box 2: Kwaku (+VER)
Me siesie ɛdan mu.
1.SG.SBJ tidy.up.PRS room inside
'*I tidy up rooms.*'

Box 3: Kwame (−EXH)
Me siesie ɛdan mu.
1.SG.SBJ tidy.up.PRS room inside
'*I tidy up rooms.*'

Box 4: Kofi
Me bisa aduane.
1.SG.SBJ request.PRS meal
'*I request a meal.*'

Experiment II (falsifier)

Box 1: Kwame
Me bisa aduane.
1.SG.SBJ request.PRS meal
'*I request a meal.*'

Box 2: Kofi (+FAL)
Me siesie ɛdan mu.
1.SG.SBJ tidy.up.PRS room inside
'*I tidy up rooms.*'

Box 3: Anan
Me hwie nsa.
1.SG.SBJ pour.PRS drink
'*I pour drinks.*'

Box 4: Kwaku (−CAN)
Me noa kosua.
1.SG.SBJ boil.PRS egg
'*I boil eggs.*'

Item 128 Anan na ɔfaa sekan.
Anan FM RPRON.take.PST knife
'*It is Anan who took a knife.*'

Experiment I (verifier)

Box 1: Kwaku
Me kyee awi.
1.SG.SBJ catch.PST thief
'*I caught a thief.*'

Box 2: Anan (+VER)
Me faa sekan.
1.SG.SBJ take.PST knife
'*I took a knife.*'

Box 3: Kofi
Me tamm obi.
1.SG.SBJ carry.PST someone
'*I carried someone.*'

Box 4: Kwame (−EXH)
Me faa sekan.
1.SG.SBJ take.PST knife
'*I took a knife.*'

Experiment II (falsifier)

Box 1: Kofi
Me kyee awi.
1.SG.SBJ catch.PST thief
'*I caught a thief.*'

Box 2: Kwame (+FAL)
Me faa sekan.
1.SG.SBJ take.PST knife
'*I took a knife.*'

Box 3: Kwaku
Me tamm obi.
1.SG.SBJ carry.PST someone
'*I carried someone.*'

Box 4: Anan (+CAN)
Me faa sekan.
1.SG.SBJ take.PST knife
'*I took a knife.*'

Item 129 Kwaku na Maanu hunahunaa no.
Kwaku FM Maanu scare.PST 3.SG.OBJ
'*It is Kwaku who Maanu scared.*'

Experiment I (verifier)

Box 1: Anan
Maanu daadaa me.
Maanu deceive.PST 1.SG.OBJ
'*Maanu deceived me.*'

Box 2: Kwaku (+VER)
Maanu hunahunaa me.
Maanu scare.PST 1.SG.OBJ
'*Maanu scared me.*'

Box 3: Kwame
Maanu pɛɛ m'asɛm.
Maanu like.PST 1.SG.POSS.matter
'*Maanu liked me.*'

Box 4: Kofi (+EXH)
Maanu teetee me.
Maanu frustrate.PST 1.SG.OBJ
'*Maanu frustrated me.*'

Experiment II (falsifier)

Box 1: Kwame
Maanu daadaa me.
Maanu deceive.PST 1.SG.OBJ
'*Maanu deceived me.*'

Box 2: Kofi (+FAL)
Maanu hunahunaa me.
Maanu scare.PST 1.SG.OBJ
'*Maanu scared me.*'

Box 3: Kwaku (+CAN)
Maanu hunahunaa me.
Maanu scare.PST 1.SG.OBJ
'*Maanu scared me.*'

Box 4: Anan
Maanu pɛɛ m'asɛm.
Maanu like.PST 1.SG.POSS.matter
'*Maanu liked me.*'

Item 131 Kofi na Dede boa no.
Kofi FM Dede help.PRS 3.SG.OBJ
'*It is Kofi who Dede helps.*'

Experiment I (verifier)

Box 1: Kwame
Dede frɛ me.
Dede call.PRS 3.SG.OBJ
'*Dede calls me.*'

Box 2: Kofi (+VER)
Dede boa me.
Dede help.PRS 3.SG.OBJ
'*Dede helps me.*'

Box 3: Anan (−EXH)
Dede boa me.
Dede help.PRS 3.SG.OBJ
'*Dede helps me.*'

Box 4: Kwaku
Dede hyɛ me nkuran.
Dede give.PRS 3.SG.OBJ encouragement
'*Dede encourages me.*'

Experiment II (falsifier)

Box 1: Anan
Dede frɛ me.
Dede call.PRS 3.SG.OBJ
'*Dede calls me.*'

Box 2: Kwaku (+FAL)
Dede boa me.
Dede help.PRS 3.SG.OBJ
'*Dede helps me.*'

Box 3: Kwame
Dede hyɛ me nkuran.
Dede give.PRS 3.SG.OBJ encouragement
'*Dede encourages me.*'

Box 4: Kofi (+CAN)
Dede boa me.
Dede help.PRS 3.SG.OBJ
'*Dede helps me.*'

Item 132 Kwame na Adwoa dii ne hwammɔ.
Kwame FM Adwoa eat.PST 3.SG.OBJ disappointment
'*It is Kwame who Adwoa disappointed.*'

Experiment I (verifier)

Box 1: Kofi
Adwoa gyegyee me.
Adwoa annoy.PST 1.SG.OBJ
'*Adwoa annoyed me.*'

Box 2: Kwame (+VER)
Adwoa dii me hwamɔ.
Adwoa eat.PST 1.SG.OBJ disappointment
'*Adwoa disappointed me.*'

Box 3: Kwaku
Adwoa suasuaa me.
Adwoa imitate.PST 1.SG.OBJ
'*Adwoa imitated me.*'

Box 4: Anan (−EXH)
Adwoa dii me hwamɔ.
Adwoa eat.PST 1.SG.OBJ disappointment
'*Adwoa disappointed me.*'

Experiment II (falsifier)

Box 1: Kwaku
Adwoa suasuaa me.
Adwoa imitate.PST 1.SG.OBJ
'*Adwoa imitated me.*'

Box 2: Anan (+FAL)
Adwoa dii me hwamɔ.
Adwoa eat.PST 1.SG.OBJ disappointment
'*Adwoa disappointed me.*'

Box 3: Kofi
Adwoa gyegyee me.
Adwoa annoy.PST 1.SG.OBJ
'*Adwoa annoyed me.*'

Box 4: Kwame (−CAN)
Adwoa teetee me.
Adwoa frustrate.PST 1.SG.OBJ
'*Adwoa frustrated me.*'

D.2.2 Removed stimuli

These stimuli contained errors and thus judgments for these trials were removed from the final analysis. The errors were as follows: Item 121 had a different verb in the auditory stimuli ('push') from the written statement ('pull'), and Item 130 had a

different part of the body in the auditory stimuli ('ear') from the written statement ('back'). As a result, these two items were invariably judged as 'false' regardless of condition.

Item 121 Kofi na Anum sum no.
Kofi FM Anum push.PRS 3.SG.OBJ
'*It is Kofi who Anum pushes.*'

Experiment I (verifier)

Box 1: Kwame
Anum hyɛ me so.
Anum oppress.PRS 1.SG.OBJ on
'*Anum oppresses me.*'

Box 2: Kofi (intended: +VER)
Anum twe me.
Anum pull.PRS 1.SG.OBJ
'*Anum pulls me.*'

Box 3: Anan
Anum sane me.
Anum untie.PRS 1.SG.OBJ
'*Anum unties me.*'

Box 4: Kwaku (intended: +EXH)
Anum suro me.
Anum fear.PRS 1.SG.OBJ
'*Anum is afraid of me.*'

Experiment II (falsifier)

Box 1: Anan
Anum sane me.
Anum untie.PRS 1.SG.OBJ
'*Anum unties me.*'

Box 2: Kwaku (intended: +FAL)
Anum twe me.
Anum pull.PRS 1.SG.OBJ
'*Anum pulls me.*'

Box 3: Kofi (−CAN)
Anum suro me.
Anum fear.PRS 1.SG.OBJ
'*Anum is afraid of me.*'

Box 4: Kwame
Anum hyɛ me so.
Anum oppress.PRS 1.SG.OBJ on
'*Anum oppresses me.*'

Item 130 Anan na Abena bɔ n'asom.
Anan FM Abena slap.PRS 1.SG.OBJ'ear (=slap)
'*It is Anan who Abena slaps.*'

Experiment I (verifier)

Box 1: Kwaku
Abena krɔkrɔ me.
Abena pamper.PRS 1.SG.OBJ
'*Abena pampers me.*'

Box 2: Anan (intended: +VER)
Abena bɔ me mu.
Abena hit.PRS 1.SG.OBJ inside
'*Abena hits my back.*'

Box 3: Kofi
Abena twerɛ me.
Abena write.PRS 1.SG.OBJ
'*Abena writes to me.*'

Box 4: Kwame (intended: +EXH)
Abena gyegye m'ani.
Abena entertain.PRS 1.SG.POSS.eyes
'*Abena makes me happy.*'

Experiment II (falsifier)