

ANTECEDENT COMPLEXITY EFFECTS ON ELLIPSIS PROCESSING

Doctoral Thesis submitted to the Faculty of Human Sciences at the
University of Potsdam in partial fulfillment of the requirements for the
degree of Doctor of Philosophy



University of Potsdam
Department of Linguistics

by

Dario Paape

2017

Supervisors:

Prof. Dr. Shravan Vasishth
Prof. Dr. Reinhold Kliegl

This work is licensed under a Creative Commons License:
Attribution 4.0 International
To view a copy of this license visit
<http://creativecommons.org/licenses/by/4.0/>

Submitted: November 27, 2017
Defended: May 8, 2018

Reviewers:

Prof. Dr. Shravan Vasishth, University of Potsdam, Germany
Prof. Dr. Brian Dillon, University of Massachusetts, Amherst

Published online at the
Institutional Repository of the University of Potsdam:
URN [urn:nbn:de:kobv:517-opus4-411689](http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-411689)
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-411689>

ABSTRACT

This dissertation explores whether the processing of ellipsis is affected by changes in the complexity of the antecedent, either due to added linguistic material or to the presence of a temporary ambiguity. Murphy (1985) hypothesized that ellipsis is resolved via a string copying procedure when the antecedent is within the same sentence, and that copying longer strings takes more time. Such an account also implies that the antecedent is copied without its structure, which in turn implies that recomputing its syntax and semantics may be necessary at the ellipsis gap. Alternatively, several accounts predict null effects of antecedent complexity, as well as no reparsing. These either involve a structure copying mechanism that is cost-free and whose finishing time is thus independent of the form of the antecedent (Frazier & Clifton, 2001), treat ellipsis as a pointer into content-addressable memory with direct access (Martin & McElree, 2008, 2009), or assume that one structure is ‘shared’ between antecedent and gap (Frazier & Clifton, 2005).

In a self-paced reading study on German sluicing, temporarily ambiguous garden-path clauses were used as antecedents, but no evidence of reparsing in the form of a slowdown at the ellipsis site was found. Instead, results suggest that antecedents which had been reanalyzed from an initially incorrect structure were easier to retrieve at the gap. This finding that can be explained within the framework of cue-based retrieval parsing (Lewis & Vasishth, 2005), where additional syntactic operations on a structure yield memory reactivation effects.

Two further self-paced reading studies on German bare argument ellipsis and English verb phrase ellipsis investigated if adding linguistic content to the antecedent would increase processing times for the ellipsis, and whether insufficiently demanding comprehension tasks may have been responsible for earlier null results (Frazier & Clifton, 2000; Martin & McElree, 2008). It has also been suggested that increased antecedent complexity should shorten rather than lengthen retrieval times by

providing more unique memory features (Hofmeister, 2011). Both experiments failed to yield reliable evidence that antecedent complexity affects ellipsis processing times in either direction, irrespectively of task demands.

Finally, two eye-tracking studies probed more deeply into the proposed reactivation-induced speedup found in the first experiment. The first study used three different kinds of French garden-path sentences as antecedents, with two of them failing to yield evidence for reactivation. Moreover, the third sentence type showed evidence suggesting that having failed to assign a structure to the antecedent leads to a slowdown at the ellipsis site, as well as regressions towards the ambiguous part of the sentence. The second eye-tracking study used the same materials as the initial self-paced reading study on German, with results showing a pattern similar to the one originally observed, with some notable differences.

Overall, the experimental results are compatible with the view that adding linguistic material to the antecedent has no or very little effect on the ease with which ellipsis is resolved, which is consistent with the predictions of cost-free copying, pointer-based approaches and structure sharing. Additionally, effects of the antecedent's parsing history on ellipsis processing may be due to reactivation, the availability of multiple representations in memory, or complete failure to retrieve a matching target.

ACKNOWLEDGEMENTS

Above all, I thank Shравan Vasishth for being the best supervisor one could wish for; for providing guidance where it was needed and otherwise putting his faith in me and my abilities. I also thank Reinhold Kliegl for sharing his deep knowledge and for being an inspiration to scientists across the disciplines.

I owe gratitude to Brian Dillon for agreeing to run one of my experiments at the University of Massachusetts, Amherst, as well as to Shayne Sloggett for overseeing the data collection process. Brian also graciously agreed to be my second reviewer.

I am grateful to Barbara Hemforth and the scientific board of the Labex EFL for hosting me at the Université Paris Diderot for three months and allowing me to run an eye-tracking experiment there, as well as for the stimulating discussions and the unparalleled hospitality I enjoyed during my stay. Furthermore, I thank Céline Pozniak and Rachel Shen for their support during the preparation and realization of the study, as well as Anne Abeillé, whose charisma, paired with insightful comments on the French materials, never failed to brighten my day.

I want to thank Bruno Nicenboim for letting me partake in his deep knowledge and for being eternally patient, no matter how many questions I asked. I am also indebted to Johanna Thieke and Daniela Merten for their help with data collection in Germany.

I thank Titus von der Malsburg, Lena Jäger, Samar Husain and everyone else I met at the Vasishth Lab for contributing to an ever inspiring environment to do research in, and for generally being the great people they are.

Finally, I want to thank Lisa, my wonderful and incomparable wife, as well as my family, just for being there, always.

CONTENTS

Abstract	ii
Acknowledgements	iv
1 Introduction	1
2 Filling the silence: Reactivation, not reconstruction	7
2.1 Abstract	7
2.2 Introduction	8
2.3 Material & Methods	13
Stimuli	13
Participants	16
Procedure	16
2.4 Results and Discussion	17
Data analysis	17
Comprehension accuracy	18
Reading times	19
Discussion	23
2.5 General Discussion	30
3 Does antecedent complexity affect ellipsis processing? An empirical investigation	33
3.1 Introduction	34
3.2 Experiment 1	38
Materials	39
Participants	40
Procedure	41
Predictions	41
Data analysis	43
Results	44
Discussion	50
A potential influence of within-sentence wrap-up	51
A possible issue of parallelism	52

	A more precise notion of complexity-based facilitation is needed	54
	Conclusion	54
3.3	Experiment 2	55
	Materials	56
	Participants	58
	Procedure	58
	Predictions	58
	Data analysis	59
	Results	60
	Discussion	62
3.4	General Discussion	66
	Conclusions	67
4	Processing of ellipsis with garden-path antecedents in French and German: Evidence from eye tracking	69
4.1	Introduction	70
4.2	Experiment 1	74
	Materials	75
	Predictions	80
	Pre-study: Off-line antecedent acceptability ratings	81
	Participants	83
	Procedure	83
	Data Analysis	84
	Results – Subject-object inversion (SOI)	85
	Results – Reduced relative clauses (RRC)	88
	Results – Triple lexical ambiguity (TLA)	92
	Discussion	97
	Exploratory analyses	102
4.3	Experiment 2	104
	Materials	104
	Predictions	106
	Participants	107
	Procedure	107
	Data Analysis	107
	Results	108
	Discussion	114
4.4	General Discussion	118
5	Conclusion	121

LIST OF FIGURES

2.1	Residual reading times for the antecedent regions, extreme values removed. Error bars represent 95% confidence intervals.	20
2.2	Residual reading times for the pre-ellipsis, ellipsis, and spillover regions, extreme values removed. Error bars represent 95% confidence intervals.	21
3.1	Experiment 1 – Mean reading times (back-transformed, on reciprocal scale) by region and condition. Error bars represent 95% confidence intervals.	45
3.2	Experiment 1 – Question response times and accuracies: caterpillar plot of means and 95% credible intervals for parameters of interest. el = elision, comp = complexity, log_rt = log response time.	46
3.3	Experiment 1 – Caterpillar plot of means and 95% credible intervals for parameters of interest, separate models fit across regions crit–2 through crit+2. el = elision, comp = complexity.	47
3.4	Experiment 1 – Caterpillar plot of means and 95% credible intervals for parameters of interest, one model fit across regions crit–2 through crit+2. <i>diff</i> parameters represent successive differences between regions, in linear order. el = elision, comp = complexity.	48
3.5	Experiment 1 – Density plots for the Bayes factor analysis. Blue = prior density, red = posterior sample density.	49
3.6	Experiment 2 – Mean reading times (back-transformed, on reciprocal scale) by region and condition. Error bars represent 95% confidence intervals.	61
3.7	Experiment 2 – Question response times and accuracies: caterpillar plot of means and 95% credible intervals for parameters of interest. comp = complexity, p_type = probe type, log_rt = log response time.	62
3.8	Experiment 2 – Caterpillar plot of means and 95% credible intervals for parameters of interest, separate models fit across regions crit–2 through crit+2. comp = complexity, p_type = probe type.	63
3.9	Experiment 2 – Caterpillar plot of means and 95% credible intervals for parameters of interest, one model fit across regions crit–2 through crit+2. <i>diff</i> parameters represent successive differences between regions, in linear order. comp = complexity, p_type = probe type.	64

3.10	Experiment 2 – Density plots for the Bayes factor analysis. Blue = prior density, red = posterior sample density.	65
4.1	Reading measures by region for SOI stimuli (Experiment 1). All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.	86
4.2	Reading measures by region for RRC stimuli (Experiment 1). All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.	91
4.3	Reading measures by region for TLA stimuli (Experiment 1). All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.	96
4.4	Reading measures by region (Experiment 2). All measures log- transformed and residualized against region length in characters; error bars show 95% intervals.	109
4.5	Comparison of effects between Paape (2016) and Experiment 2. . .	116

LIST OF TABLES

2.1	Untrimmed raw mean reading times in milliseconds by condition for antecedent, ellipsis and spillover regions, standard errors in parantheses.	18
2.2	Coefficient estimates, standard errors and t values for the linear mixed-effects models fit to reciprocal reading times at the indicated regions of interest.	22
2.3	Coefficient estimates, standard errors and t values for the linear mixed-effects model fit to reciprocal reading times at region wh-1.	29
4.1	Pre-study: Mean acceptability ratings and standard errors (in parantheses) by sentence type and condition.	82
4.2	Experiment 1: Results for question responses accuracy and response times (SOI stimuli). can = canonicity, el = elision.	87
4.3	Experiment 1: Results for first-pass reading times (SOI stimuli). can = canonicity, el = elision.	88
4.4	Experiment 1: Results for regression-path durations (SOI stimuli). can = canonicity, el = elision.	89
4.5	Experiment 1: Results for total reading times (SOI stimuli). can = canonicity, el = elision.	90
4.6	Experiment 1: Results for question response accuracy and response times (RRC stimuli). amb = ambiguity, el = elision.	92
4.7	Experiment 1: Results for first-pass reading times (RRC stimuli). amb = ambiguity, el = elision.	93
4.8	Experiment 1: Results for regression-path durations (RRC stimuli). amb = ambiguity, el = elision.	94
4.9	Experiment 1: Results for total reading times (RRC stimuli). amb = ambiguity, el = elision.	95
4.10	Experiment 1: Results for question response accuracy and response times (TLA stimuli). amb = ambiguity, el = elision.	97
4.11	Experiment 1: Results for first-pass reading times (TLA stimuli). amb = ambiguity, el = elision.	98
4.12	Experiment 1: Results for regression-path durations (TLA stimuli). amb = ambiguity, el = elision.	99
4.13	Experiment 1: Results for total reading times (TLA stimuli). amb = ambiguity, el = elision.	100

4.14	Experiment 2: Results for question response accuracy and response times. case = case marking, ord = word order.	108
4.15	Experiment 2: Results for question response accuracy by probe type. case = case marking, ord = word order, pt = probe type.	110
4.16	Experiment 2: Results for first-pass reading times. case = case marking, ord = word order.	111
4.17	Experiment 2: Results for regression-path durations. case = case marking, ord = word order.	112
4.18	Experiment 2: Results for total reading times. case = case marking, ord = word order.	113

CHAPTER 1

INTRODUCTION

The aim of this thesis is to investigate the processing of nothing. Nothing, in this case, specifically means the absence of words. An absence, however, that conveys meaning, and the mechanism by which this meaning comes to be is the topic of the present work.

Part of the enduring appeal of ellipsis as a research subject, to me, appears to be that it is not an observable thing. What is observed is the *product* of ellipsis, namely, a lack of words where words should be. In (1), nothing apart from the fact that the sentence is grammatical even though it is lacking a verb in the second conjunct tells the reader that ellipsis has applied. Yet, figuring out that the word *knows* is missing after *Vivian* does not require conscious effort.

- (1) Kyle knows Ben and Vivian Sandra.

Ellipsis, though it possesses anaphoric properties, is – at least on the surface – very unlike, say, a personal pronoun such as *him*. The latter is a lexical item that has the built-in property of being a stand-in for another expression. The former needs to be inferred from the fact that no structure for (1) can be derived without assuming that otherwise mandatory parts of the utterance have been suppressed. In a way, ellipsis is thus more similar to syntactic traces, which are assumed to occupy phonologically empty positions from which an expression has been moved (Chomsky, 1976). Just like traces are assumed to be generated by a rule of movement, an ellipsis gap can be thought of having been created by a deletion rule (cf. Merchant, 2001; Ross, 1967), possibly triggered by a desire to compress the discourse and avoid redundancy.

However, it has repeatedly been pointed out that this view is too simplistic, as ellipsis is apparently licensed in many contexts for which there is no ‘un-deletion’ such that literally repeating (part of) the antecedent would yield a grammatical sentence (e.g. Kehler, 2000):

- (2) a. The sticker should have been removed but seemingly nobody did.
b. The sticker should have been removed but seemingly nobody did [remove the sticker].

This quirky property of ellipsis has led some scholars to treat it as a semantic or discourse phenomenon that does not operate on surface strings or syntactic phrases (Hardt, 1993; Kehler, 2000). Then again, certain kinds of ellipsis delete a whole clause and leave behind a case-marked *wh*-pronoun, which is mysterious if the suppressed elements do not participate in the syntactic derivation (Merchant, 2001). Other syntactic constraints, such as Condition A of the Binding Theory, also seem to be observed in ellipsis contexts (Kennedy, 2003). Given observations such as these, Merchant (2001, 2013) argues that syntactic structure must be present at the ellipsis site, and that the deletion process must be sensitive to this structure.

To this day, no final consensus has emerged in the theoretical literature as to what level of linguistic representation ellipsis operates on, or whether to even characterize ellipsis as a process, rather than as an (invisible) entity, or as a combination of both. Studies investigating the on-line processing of ellipsis have also not yielded an unambiguous picture regarding these questions, as conflicting results have been reported over the years (see review by Phillips & Parker, 2014). For instance, it has been claimed that ellipsis does not induce syntactic priming (Cai, Pickering, & Sturt, 2013), which would support a syntax-based view, but a more recent study reports such priming effects (Xiang, Grove, & Merchant, 2014).

In the present work, I neither subscribe to any particular view regarding the kind of antecedent representations (semantic, syntactic, discourse-level) that are accessed during ellipsis processing, nor regarding the status of the ellipsis site itself with regard to the presence or absence of syntactic structure. Rather, I focus on the more basic distinction between structured and unstructured antecedent representations in a general sense, between simple and complex antecedents, as well as on effects of the antecedent's processing history on the processing of the ellipsis.

My agnosticism regarding the issue of representation may strike some readers as unsatisfactory, but I will attempt some further justification. Compare (3a,b), both taken from Murphy (1985), as well as (3c), first mentioned by Montalbetti (1984).

- (3) a. We can't tell you who banks at Morgan. But we can tell you why.
b. A: Was the cake taken by you? B: * No, Sandy did.
c. More people have been to Russia than I have.

Murphy (1985, p. 291) notes that (3a) appeared in an advertisement even though “[t]he anaphor [...] has no correct linguistic antecedent”. He also judges (3b) as ungrammatical, which contrasts with the (arguable) acceptability of (2a) above. At the very least, these examples demonstrate that the acceptability of ellipsis is highly dependent on subtle contextual changes, and that there is potentially a large amount of variability in speakers’ judgments, which makes it difficult to argue for or against any one theory based on isolated pieces of evidence.

An even more extreme example is (3c), a so-called ‘dead end’ sentence for which it should be impossible to derive a structure. It is not clear what the intended meaning of the utterance is – that more people besides me have been to Russia? That there are people who have visited Russia more often than I have? Yet, most native speakers of English experience a persistent illusion of grammaticality when presented with (3c) (Phillips, Wagers, & Lau, 2011). This illusion, in turn, is argued by Fulst and Phillips (2004) to be due to ‘repair by ellipsis’, where the corresponding unelided form would be (4).

(4) More people have been to Russia than I have been to Russia.

The troubling observation is that (4) is just as ungrammatical/nonsensical as (3c), but in (4) readers appear to notice the strangeness more readily, as if by assuming that ellipsis has applied in (3c) “the syntactic problem [...] is somehow eliminated from detection” (Wellwood, Pancheva, Hacquard, & Phillips, 2017, p. 39).¹

To my mind, such examples show that it is well advised to tread lightly around ellipsis, given that it is apparently able to warp a speaker’s sense of what constitutes an acceptable utterance. For this reason, I have chosen the somewhat less audacious approach of investigating the processing difficulty of ellipsis in (hopefully) uncontroversially grammatical constructions without trying to determine precisely which kinds of representations the processes involved operate on. For example, whenever the term ‘syntactic reanalysis’ is used in the following chapters, note that this usually entails semantic reanalysis, as well as possible reanalysis of discourse representations. Likewise, the notion of complexity I employ refers to syntactic, semantic and discourse-level complexity alike.

The reader may also notice that I do not dwell overly long on differences between particular types of ellipsis, such as clausal versus verb phrase ellipsis, which undoubtedly and relevantly exist. However, I made an attempt at abstracting away from such differences and focus

¹While the effect was not replicated in the experiments of Wellwood et al. (2017), their unelided control sentences avoided verbatim repetition of the second clause by substituting *Canada* for *Russia*, which essentially makes the conditions uncomparable, as it doesn’t directly address the question of whether ellipsis increases acceptability anymore.

instead on the commonalities, namely the requirement of accessing an antecedent in memory and integrating it at the gap. At this point, it should be noted that devising a theory that accounts for the shared properties of all ellipsis varieties in a principled way would be preferable to having a separate theory for each.

The thesis is composed of three journal articles. The empirical question posed by the first article is whether processing ellipsis requires the reader to reparse the antecedent, where ‘parse’ should be taken to mean ‘take the steps necessary for successful interpretation’, which implies ‘assign structure’. Sentences in which the antecedent is a garden-path structure are used as a test case. The second article asks the closely related question of whether processing ellipsis becomes more time-consuming when the antecedent contains more information. Furthermore, it investigates whether such complexity effects are affected by task demands, specifically by how deeply comprehension is probed. Drawing on earlier results, the third article explores whether ellipsis becomes easier to process if the antecedent was structurally reanalyzed at an earlier point in the sentence.

Using the self-paced reading method, the first research article, as presented in Chapter 2, investigated whether ellipsis takes longer to process when the antecedent is a temporarily ambiguous garden-path sentence. In garden-path sentences, the parser adopts an erroneous syntactic analysis for the ambiguous part of the utterance before being forced to reanalyze by disambiguating material. It has been suggested repeatedly in the processing literature that the resolution of ellipsis requires copying of antecedent material into the ellipsis gap, at least when the antecedent is ‘close’ to the ellipsis site (Frazier & Clifton, 2001; Murphy, 1985). Meanwhile, other accounts uphold that ellipsis is no more than a pointer to the antecedent information that has been kept in working memory (Martin & McElree, 2008, 2009). A third account claims that antecedent and ellipsis gap ‘share’ one and the same structure (Frazier & Clifton, 2005). The main difference between the first type of account on the one hand and the second and third approaches on the other is that copying implies duplication of information (cf. Martin & McElree, 2008). Furthermore, Murphy (1985) explicitly assumes that what is transcribed is a word string, which implies that the resulting copy at the gap will initially be unstructured. Given this assumption, it is possible that despite having resolved a temporary ambiguity in the antecedent, the parser will also experience some difficulty at the ellipsis site, as only words and no structure are copied over. The experiment, carried out in German with a well-known type of subject-object ambiguity (Hemforth, 1993; Meng & Bader, 2000), found no evidence that reparsing at the ellipsis site was necessary. Instead, results showed an apparent mismatch penalty for non-canonical antecedents at the ellipsis site, accompanied by an unexpected speedup for antecedents with a garden-path structure. Both results can be explained within the cue-based retrieval parsing framework of Lewis and Vasishth (2005) if the

quality of the antecedent's match with the retrieval cues as well as higher memory activation of reanalyzed antecedents are assumed to influence processing speed at the ellipsis site.

The second article, which is reproduced in Chapter 3, presents two further self-paced reading experiments, carried out in German and English, respectively. Experiment 1 was designed to investigate a further prediction of Murphy's (1985) string copying account, namely that adding additional material to the antecedent should increase processing times at the ellipsis site. Murphy (1985) found evidence for such an effect, but later studies report null results (Frazier & Clifton, 2001; Martin & McElree, 2008). Furthermore, under the account of Hofmeister (2011), more complex memory representations should become *easier* instead of more difficult to retrieve, as additional features help create a more distinctive trace. Results show evidence in favor of a null effect of antecedent complexity at the critical ellipsis region. This observation is qualified, however, by the possibility of a very small slowdown that was not detectable given the study's power, as well as by a possible interaction with a sentence-level speedup (Demberg & Keller, 2008; Ferreira & Henderson, 1993) that would have canceled out the complexity effect. Based on findings indicating that task demands influence on-line sentence processing (Foertsch & Gernsbacher, 1994; Swets, Desmet, Clifton, & Ferreira, 2008), Experiment 2 aimed to probe into a possible interaction between participants' depth of processing, manipulated through comprehension questions, and antecedent complexity. Again, results do not show evidence that the antecedent's complexity affects the processing of the ellipsis, even when comprehension query its interpretation directly, as opposed to requiring only superficial memory of the sentence. Overall, the findings presented in Chapter 3 are most consistent with the claims of pointer-based approaches to ellipsis processing, which predict no effects of antecedent complexity.

In the third article of this thesis, presented in Chapter 4, the post-hoc explanation of the surprising finding of Chapter 2, namely that reactivation of the antecedent's memory trace through reanalysis leads to faster processing of the ellipsis, is subjected to further empirical scrutiny. Experiment 1 is an eye-tracking study on three different types of garden-path structures in French that to our knowledge had not been systematically investigated before: sentences with OVS word order, comparable to the stimuli used in Chapter 2's experiment, sentences containing reduced relative clauses and sentences featuring a series of lexical ambiguities. While all three stimulus types created on-line garden-path effects, only those involving lexical ambiguities showed an interaction with elision later on. Furthermore, the interaction went in the opposite direction compared to the initial prediction, such that ellipsis created more processing difficulty when the antecedent was a garden-path structure. This effect is argued to be due to participants' occasional use of 'good-enough' processing (e.g. Karimi & Ferreira, 2016), where the antecedent is left unanalyzed and

becomes irretrievable as a result. Experiment 2 was an attempt to replicate the main finding of Chapter 2 using eye tracking. On the same materials used in the original experiment, first-pass and total reading times showed the critical interaction, albeit at the cost of comprehension accuracy. The inconsistent results for French and German, as well as the source of the differences between sentence types in the French study, are discussed.

Chapter 5 again summarizes the major results of the present work, and makes an attempt to fit the various theoretical and empirical implications into a unified picture.

CHAPTER 2

FILLING THE SILENCE: REACTIVATION, NOT RECONSTRUCTION

Dario L. J. F. Paape

FRONTIERS IN PSYCHOLOGY, VOLUME 6, ISSUE 27

Abstract

In a self-paced reading experiment, we investigated the processing of sluicing constructions ('sluices') whose antecedent contained a known garden-path structure in German. Results showed decreased processing times for sluices with garden-path antecedents as well as a disadvantage for antecedents with non-canonical word order downstream from the ellipsis site. A post-hoc analysis showed the garden-path advantage also to be present in the region right before the ellipsis site. While no existing account of ellipsis processing explicitly predicted the results, we argue that they are best captured by combining a local antecedent mismatch effect with memory trace reactivation through reanalysis.

Introduction

Besides verb-phrase ellipsis, sluicing (Ross, 1969) is probably the most-studied ellipsis variety in both theoretical linguistics (e.g., Chung, Ladusaw, & McCloskey, 1995; Merchant, 2001; Potsdam, 2007) and psycholinguistics (e.g., M. W. Dickey & Bungler, 2011; Poirier, Wolfinger, Spellman, & Shapiro, 2010; Yoshida, Dickey, & Sturt, 2013). In sluicing, an entire clause is left out and a *wh*-element remains behind, as in (5).

- (5) John saw Mary, but I don't remember when ____.
 ____ = John saw Mary

Sluicing is anaphoric: to interpret (5), the semantics of the antecedent (*John saw Mary*) must somehow be inserted into the gap behind the word *when* to derive the meaning *I don't remember when John saw Mary*. We write 'meaning' because deriving an interpretation is the fundamental goal of sentence processing, not because it is necessarily clear that the relevant representation of the antecedent is semantic in nature. There is an ongoing debate as to whether syntactic structure is also present at ellipsis sites (cf. Cai, Pickering, & Sturt, 2013, and references therein), or whether one should adopt a more discourse-centered approach to the gap-filling process (e.g., Hardt, 1993; Kehler, 2000). Since the evidence to date, at least in our view, does not unequivocally favor any of these views, we will not take a stance with regard to the representation question. We will, however, use syntactic terminology throughout the article for ease of reference.

Even with the question of *what* is inserted into the gap set aside, another point of debate has been *how* it ends up there. Ross (1967) was perhaps the first to explicitly propose a deletion approach to ellipsis (in this case, verb-phrase ellipsis): the missing bit of structure is assumed to be underlyingly present, but its phonological representation is erased under identity with the antecedent.¹ Under the approach taken by Williams (1977), ellipsis involves copying. Like Ross (1967), Williams assumes invisible syntax at the gap, but the terminal symbols of this structure are null elements (Wasow, 1972). The ellipsis is interpreted by copying the terminals (that is, words) from the antecedent to the appropriate positions within the gap.

From a processing perspective, it is not enough to claim that the syntax is there in the silence: the processor must have some way of creating it. A reader of (5) would have to first infer that deletion has applied, then identify the antecedent and finally reconstruct it at the

¹There is no condition of strict identity, however, as several kinds of mismatch can be observed, as in *The car was supposed to be washed but nobody did ~~wash the car~~* (e.g., Kertz, 2000; Merchant, 2013, 2016).

gap. The main aim of the current study is to investigate how this ‘reconstruction’ is to be conceived of: does the parser rebuild the antecedent’s structure at the ellipsis site, or does it come to be there by virtue of some other mechanism?

One might think of dispensing with the idea of invisible structure altogether. The approach of Hardt (1993) is explicitly non-syntactic in nature and treats ellipsis as an unstructured proform that refers to a stored meaning in a discourse model. The notion of copying does not enter into the picture; ellipsis acts rather like a pointer or a hyperlink into memory than as an entity of its own. This conception can be related to the processing of other types of anaphors: It is not commonly assumed that in a sentence such as *The man from England drank tea, but he didn’t drink coffee*, the pronoun *he* will contain the syntactic structure of the NP *the man from England* at any level of representation. Instead, an identity of *reference* between the two expressions seems to obtain (cf. Grinder & Postal, 1971, p. 269).

Note that the opposition between copying and the ‘memory pointer’ approach is orthogonal to that between syntactic and semantic/discourse representations (cf. Phillips & Parker, 2014). Semantic representations could also be copied, just as syntactic representations could be pointed to. The processing literature has focused mainly on the copying/pointing dichotomy, even though some studies have also tested whether there is syntactic priming from ellipsis sites, with mixed results (Cai, Pickering, & Sturt, 2013; Xiang, Grove, & Merchant, 2014). Murphy (1985) appears to have been the first to systematically look for effects of antecedent length on reading times for elliptical clauses, in this case the sentence *Later, his uncle did too* in (6).

- (6) a. Jimmy swept the floor. Later, his uncle did too.
b. Jimmy swept the tile floor behind the chairs free of hair and cigarettes. Later, his uncle did too.

Despite being concerned with verb-phrase ellipsis, we assume that this study is informative with regard to sluicing as well, since the most parsimonious hypothesis would be that all types of ellipsis are processed in the same way. The reasoning behind Murphy’s manipulation was that “[l]onger antecedents would be expected to affect a copying process, since the longer the string that must be copied onto the anaphor, the longer it should take to understand the anaphor” (p. 293). If there was no copying, so the argument goes, then reading times for the second sentence should not differ between (6a,b). Murphy found that reading times for the elliptical sentence were increased by about 260 ms when the antecedent was long rather

than short. Interestingly, this difference disappeared when another sentence was inserted between antecedent and ellipsis.²

The system Murphy proposes is one in which there are two processes, namely copying and discourse-based ‘plausible reasoning’, which operate in parallel, with the process that finishes first supplying the antecedent. When the antecedent is far away, the speed and/or availability of copying suffers and readers fall back on plausible reasoning, which by assumption is not influenced by complexity effects. (Tanenhaus & Carlson, 1990, p. 261) remain unconvinced by Murphy’s (1985) evidence for copying, arguing that the length manipulation “also introduced potential scope and attachment ambiguities”.³ The authors favor a pointer-based approach, while allowing for the possibility that there are both a syntax- and a discourse-based process at work.

Two additional important findings come from an experiment by Frazier and Clifton (2000) and a series of experiments by Martin and McElree (2008), all on verb-phrase ellipsis.

- (7) Frazier and Clifton (2000), Experiment 1 B
 - a. Sarah left her boyfriend last May. Tina did too.
 - b. Sarah got up the courage to leave her boyfriend last May. Tina did too.

- (8) Martin and McElree (2008), Experiment 3
 - a. The history professor understood Roman mythology, . . .
 - b. The history professor understood Rome’s swift and brutal destruction of Carthage, . . .

...but the principal was displeased to learn that the over-worked students attending summer session did not.

Frazier and Clifton’s study used self-paced reading and found no difference in reading times between (7a,b) for the sentence *Tina did too*. Martin and McElree’s Experiment 3, which used sentences such as (8a,b), employed a speed-accuracy trade-off paradigm with end-of-sentence acceptability judgments. No effect of antecedent complexity on processing times was

²Murphy was concerned that the observed complexity effect was simply due to processing spillover from the antecedent sentence into the ellipsis sentence, but the intervening sentence did not show any effects either.

³It is not obvious which ambiguities the authors are referring to, or how they would impact processing under an approach without copying. It should be pointed out, however, that interpreting the ellipsis with the long antecedent in (6) requires an additional assumption, namely that the floor became dirty again between the first and the second sweeping.

observed in this study and two further experiments, which the authors interpret as evidence for a pointer-based approach.

Here is where terminology becomes an issue, as Frazier and Clifton (2001) explain their earlier results by means of a mechanism called Copy α . Copy α becomes available when the scope of an ellipsis can be uniquely identified and serves as a shortcut to syntactic structure: instead of being built step-by-step, which would be computationally costly, the silent syntax is copied from the antecedent. As this process is assumed to be ‘cost-free’, the complexity of the copied structure has no influence on processing time. Frazier and Clifton’s use of the copying metaphor is not very intuitive (cf. Martin & McElree, 2008, p. 882f.), as it should take more time to copy a larger amount of information, in concordance with Murphy’s (1985) prediction.⁴ Indeed, (Frazier & Clifton, 2001, p. 17) themselves explain that a pointer would be a possible implementation of Copy α and in a later paper (Frazier & Clifton, 2005) describe Copy α as equivalent to ‘sharing’ one structure between antecedent and ellipsis (cf. also Murguia, 2004). We will thus treat pointer-based approaches, Copy α and ‘sharing’ as variants of one and the same idea, namely that the antecedent’s structure is available in memory and can be retrieved from there as-is, without any additional costly computations.

(Phillips & Parker, 2014, p. 91) make note of several methodological problems in both of the above studies. Frazier and Clifton’s (2000) experiment used only a small number of experimental items, all of which had the ellipsis at the very end of a sentence, where wrap-up effects might mask an influence of antecedent complexity. Additionally, comprehension questions were not asked after every trial and never targeted the interpretation of the ellipsis.⁵ The ungrammatical sentences in Martin and McElree’s (2008) study replaced the subject of the elliptical clause by an inanimate NP (*the overly worn books*), thus making the judgments fairly easy and possibly leading subjects to engage in superficial processing. Given these concerns, Phillips and Parker judge the results to be inconclusive, but also point out that it would be difficult to design an experiment that would provide convincing evidence for or against complexity effects.

Given this state of affairs, we think it worthwhile to look back at Frazier and Clifton’s (2001) distinction between a syntactic structure that is computed step-by-step and one that is

⁴A possible analogy would be the copying of a file from one location on a hard drive to another, which becomes more time-consuming as file size increases.

⁵An anonymous reviewer suggests that this might have been in order not to risk making the subjects aware of the experimental manipulation. While this is a fair point, it has been shown that subjects adapt their processing strategy to task demands, trying to minimize effort through underspecification (e.g., Foertsch & Gernsbacher, 1994; Swets, Desmet, Clifton, & Ferreira, 2008). If one intends to investigate ‘deep’ processing, we believe that the latter risk outweighs the former, being aware that the opposite stance is equally tenable.

retrieved from memory. What happens when the antecedent is structured in a way that is known to fool the ‘normal’ incremental parsing mechanism, that is, if it contains a garden path? Assuming a serial parsing architecture, recovering from a syntactic misanalysis involves reanalyzing the ambiguous region and assigning the same structure that would be computed for an unambiguous control sentence. Since the final memory representations for ambiguous and unambiguous sentences are the same, pointer-based approaches and Copy α would predict that there should be no difference in processing times at the ellipsis site. If, on the other hand, ellipsis is not resolved by linking the gap to a complete structure in memory, different scenarios are possible. One would be that the antecedent is accessed in memory as a word string, and that syntax and semantics are assigned to this string in the usual way, that is, incrementally. However, as verbatim memory is known to be highly fallible even in recognition tests (Sachs, 1967; Murphy & Shapiro, 1994), it may be unrealistic to assume that strings are *recalled* literally for ellipsis processing. The account of Kim, Kobele, Runner, and Hale (2011) proposes that not the words themselves but their features are accessed by the parser at the ellipsis site, and that “derivations in an initial conjunct [are allowed] to do double-duty in a second conjunct” (p. 346). Their account states that “once [...] an appropriate antecedent is found, [its derivation] becomes available to the parser, just as if it were located at the elision point in the input string” (ibid.), essentially claiming that the derivation is carried out twice. Now, if the sentence processor has no way of ‘remembering’ that it was garden-pathed by the antecedent, there is a chance that it will be garden-pathed again at the ellipsis site.

A model that is, in principle, compatible with both the pointer/sharing approach and the ‘reparsing’ account is the cue-based retrieval parser of Lewis and Vasishth (2005). In this model, syntactic phrases are stored in working memory as chunks than can be retrieved if needed. For complex phrases, both the phrase itself and its constituent parts, such as the subject of a verb phrase, are stored, along with their grammatical relations. When an ellipsis site is encountered, the parser would thus have the opportunity to retrieve either the whole antecedent as one chunk, as under a pointer-based account, or to retrieve whatever chunks are contained within the antecedent and build a new structure, as under the ‘reparsing’ view. The latter possibility may become especially attractive in cases of antecedent-ellipsis mismatch, where a strict isomorphism condition cannot be upheld (e.g., Merchant, 2001). As in the case of Kim et al. (2011) chunks are conceived of as feature bundles and thus no verbatim memory of the antecedent is required for retrieval. In fact, both Kim et al. (2011) and Lewis and Vasishth (2005) explicitly assume that the linear order of constituents is not represented in the syntax.

The ‘parse twice’ approach might seem counterintuitive, but is in fact no less parsimonious than Frazier and Clifton’s Copy α , given that it needs no special machinery besides access to grammatical features inside the antecedent structure. One would not expect the garden-path

effect at the ellipsis site to be of the same strength as the one observed for the antecedent, just as one would not expect the reading time for *when* in (5) to be equal to that of *John saw Mary*. Several steps involved in lexical access can be omitted during ellipsis processing. Simner and Smyth (1999) suggest that instead of using lexemes, ellipsis targets word lemmas, which would be compatible with the ‘feature bundle’ view described above. Additionally, ellipsis normally occurs in environments that feature a high amount of syntactic parallelism. If a parallel structure is expected, the relevant routines may be activated beforehand or at least be assigned a higher rank when the parser decides which structure to build at the ellipsis site, which can be seen as an instance of syntactic priming (Dubey, Keller, & Sturt, 2008; M. W. Dickey & Bunker, 2011). Given this assumption, however, it might be that in case of a garden path the preferred but incorrect structure will feature into the calculation, making the ellipsis more difficult to process than in cases where the antecedent’s structure is unambiguous. While Arai, Nakamura, and Mazuka (2014) found evidence that resolving an ambiguity in a prime sentence makes processing of the same ambiguity in the target sentence easier when the same verb is repeated (see also Branigan, Pickering, & McLean, 2005), it is unclear whether ellipsis constitutes ‘repetition’.

In our experiment, we used a known garden-path structure in German to test the – equivalent – predictions of pointer- and sharing-based approaches against those of a reconstruction-based approach of ellipsis processing. The former two predict that garden-pathing within the antecedent clause should have no effect at the ellipsis site while the latter predicts that the pattern observed at the point of disambiguation will reappear, although the effect size may be significantly smaller. To anticipate the results, we found an unpredicted pattern that was inconsistent with a reconstruction approach, but compatible with pointer- and sharing-based accounts if additional assumptions are made.

Material & Methods

Stimuli

It is known that German readers prefer to assign a subject interpretation to a sentence-initial NP that is ambiguous between a subject and an object reading, which results in a garden path when it is disambiguated towards an object role (cf. also Hemforth, 1993, among others). Different explanations for the subject preference have been proposed. For instance, Gorrell’s (1996) approach assumes that the parser favors structural simplicity; under his analysis,

deriving an OVS structure requires more movement operations (and thus more traces) than deriving an SVO structure, where the object presumably remains in the position at which it is base-generated. Schlesewsky, Fanselow, Kliegl, and Krems (2000) consider the possibility that the subject preference is due to a frequency-based ‘tuning’ effect (e.g., Mitchell, Cuetos, Corley, & Brysbaert, 1995), reporting over 90% nominative-initial main clauses in a corpus study. Still other possibilities are that subject-first is a default parsing assumption, as has been proposed for English (e.g., Bever, 1970; Fishbein & Harris, 2014; Grodzinsky, 1986). If one follows the current standard analysis of German clause structure, where S(O)V word order is assumed to be basic and all other word orders are derived through movement (e.g., Schwartz & Vikner, 1996/2007), the reanalysis of an object-initial structure will minimally involve removing co-indexation between an assumed trace position for the subject and the initial noun phrase, as well as postulating a trace position for an object.

The garden-path effect incurred by the non-canonical structure is stronger when disambiguation is achieved through agreement on the finite verb rather than through case marking on another NP (Meng & Bader, 2000). As shown in (9), we used indefinite NPs instead of the wh-marked NPs employed by Meng and Bader. Case marking on the *sympathizer* NP is either ambiguous (9a/b) or unambiguous (9c/d). The auxiliary *hatte(n)*, ‘had’, agrees either with the singular *sympathizer* or with the plural *rebels* NP, thereby signaling either OVS (9a/c) or SVO word order (9b/d). The result is a 2×2 design with the factors word order and case marking. Diamonds indicate the boundaries of presentation regions in the experiment, subscripts indicate region coding for the statistical analysis.

(9) a. **Ambiguous / OVS**

Eine Sympathisantin

A sympathizer.fem.**nom/acc**

der Opposition_{np1} ◇ hatten_{aux} ◇ die Rebellen_{np2} ◇ ...
of the opposition **had.pl** the rebels.nom/acc

b. **Ambiguous / SVO**

Eine Sympathisantin

A sympathizer.fem.**nom/acc**

der Opposition_{np1} ◇ hatte_{aux} ◇ die Rebellen_{np2} ◇ ...
of the opposition **had.sg** the rebels.nom/acc

c. **Unambiguous / OVS**

Einen Sympathisanten

A sympathizer.masc.**acc**

der Opposition_{np1} ◇ hatten_{aux} ◇ die Rebellen_{np2} ◇ ...
of the opposition **had.pl** the rebels.nom/acc

d. **Unambiguous / SVO**

Ein Sympathisant

A sympathizer.masc.nom

der Opposition_{np1} ◇ hatte_{aux} ◇ die Rebellen_{np2} ◇ ...
of the opposition **had.sg** the rebels.nom/acc

... laut einem Bericht_{adj} ◇ maßgeblich unterstützt_{vp}
according to a report decisively supported

◇ aber ◇ die Regierung ◇ konnte ◇ nicht ◇ nachweisen_{wh-1} ◇ wie_{wh}
but the government could not substantiate **how**

◇ so sehr_{wh+1} ◇ sich_{wh+2} ◇ die Untersuchungskommission_{wh+3} ◇ auch ◇ bemühte.
so greatly itself the investigative commission too struggled

‘The rebels had supported a sympathizer (OVS, a/c) / A sympathizer had supported the rebels (SVO, b/d), but the government could not substantiate how, no matter how hard the investigative commission tried.’

The antecedent clause ends at *unterstützt*, ‘supported’. It is conjoined with a second clause by *aber*, ‘but’, which contains a sluicing site (or ‘sluice’) at *wie*, ‘how’. All wh-phrases in the experiment were ‘sprouted’ (Chung et al., 1995), that is, they had no explicit correlate in the antecedent. We only used adjunct wh-phrases since argument wh-phrases are case-marked in German, which would have introduced a potential confound. The other wh-phrases used were several expressions meaning ‘why’ (*warum*, *weshalb*, *wieso*), *wo*, ‘where’, *wann*, ‘when’, *womit*, ‘with what’, *wozu* ‘to what (end)’, and *wobei*, ‘at what’ (combined with the verb *unterstützen*, ‘to support’). The part of the sentence following the sluicing site was intended as a spillover region. We could have used only conditions (9a) and (9c) to look for an effect of reanalysis, but decided to also include (9b) and (9d) as control conditions since otherwise reanalysis would be completely confounded with the gender of the initial NP. Additionally, even though condition (9b) is initially ambiguous, there should be no reanalysis as readers will assume SVO order by default (cf. Meng & Bader, 2000); we can thus control for temporarily ambiguous antecedents being processed differently from unambiguous ones. Thirty-two sentences were created according to this schema for use in the experiment. A complete list of the experimental materials is given in Appendix I. The stimuli were combined with 96 filler sentences featuring various constructions.

We expected a garden-path effect to occur at the auxiliary of the antecedent clause in the form of a word order × case marking interaction. Meng and Bader (2000) observed longer reaction times in a grammaticality judgment task for OVS than for SVO sentences, indicating that OVS order is overall more difficult to process. In (9a), however, the *sympathizer* NP presumably has to be reanalyzed from subject to object, which should further increase processing

time. If ellipsis acts as a pointer into memory, no interaction between the experimental factors should appear at *wie*, ‘how’, as neither the scope of the ellipsis nor the availability of a completely analyzed antecedent structure vary between conditions. If, however, the syntax of the ellipsis site has to be constructed by normal parsing routines, the garden-path effect should reappear at this position, though most likely with reduced magnitude.

We had no specific predictions as to possible effects of OVS versus SVO word order at the ellipsis site, but a post-hoc hypothesis will be developed in the discussion section. A complication concerning the predictions of both accounts that did not become apparent to us until after the experiment is that inserting a verb-second antecedent into the ellipsis site verbatim is impossible in our stimuli, as German subordinate clauses are generally required to be verb-final. The predictions outlined above are valid for well-formed antecedents, but should pertain to mismatched antecedents as well if certain additional assumptions are made, as will be explained shortly.

Participants

Sixty students from the University of Potsdam were recruited for the study. All subjects were native speakers of German and were either paid 6 € or received course credit for the participation. Informed consent was obtained from all participants prior to testing.

Procedure

The sentences were presented using the moving window self-paced reading technique (Just, Carpenter, & Woolley, 1982), which was implemented using the *Linger* software (Rohde, 2003). Participants sat in front of a PC in a quiet room and were instructed to read silently and at their own pace. Sentences were presented in 20 pt Courier New font according to a latin square procedure. At the beginning of each trial, all characters were masked with underscores. Participants completed two practice trials before the experiment proper. The order of fillers and experimental sentences was randomized at runtime. Each trial was followed by a comprehension test which took one of two forms: either a statement about the preceding sentence had to be judged as true or false, or a gap in a statement had to be filled by selecting one out of four options. Some test statements targeted the argument structure of the antecedent (*Rebels had supported a sympathizer of the opposition. [Yes/No]*), while others targeted other kinds of information from the sentence. The ratio of true to false statements

for the judgment test was balanced. For a subset of fill-in-the-gap statements appearing after experimental sentences, participants had to supply the critical wh-pronoun.⁶

Results and Discussion

Data analysis

After 15 participants had completed the experiment, it was noticed that three experimental items contained a typographical error in one condition each. The errors were removed and data from the corresponding trials were excluded from the statistical analysis. The remaining data were analyzed using the R software environment (R Core Team, 2016) by fitting linear mixed-effects models to individual regions of interest with the `lme4` package (Bates, Mächler, Bolker, & Walker, 2015). The models included varying intercepts and slopes by subjects and by items. The code and data will be released with the publication of this paper. When the estimate for a slope adjustment was zero, the random effect was dropped from the model, along with any associated higher-order effects. When a model failed to converge, random effects were removed, starting with the effect that accounted for the smallest amount of variance, until convergence was obtained. Sum contrasts were defined for the experimental factors word order and case marking and entered into the models as fixed effects. For word order, the OVS conditions were coded as 1 and the SVO conditions as -1 , respectively. For case marking, the ambiguous conditions were coded as 1 and the unambiguous conditions as -1 . Since processing spillover is a known concern in self-paced reading, the reading time for the immediately preceding region was also entered into all models after being appropriately transformed (see below) and subsequently centered. The addition of this parameter improved model fit for all regions of interest,⁷ but the method is by no means guaranteed to eliminate spillover entirely, for instance if subjects postpone processing and keep ‘tapping’ the button at fixed time intervals (Witzel, Witzel, & Forster, 2012).

An underlying assumption in linear modeling is that the residuals are approximately normally distributed. As this was not the case when raw reading times were used as the dependent variable, we applied the Box-Cox procedure (Box & Cox, 1964; Venables & Ripley, 2002), which

⁶Though the specific example in (9) was not accompanied by this kind of test, a possible fill-the-gap statement could have been *The government could not substantiate ___ rebels had supported a sympathizer of the opposition. [why/how/when/if]*.

⁷Improvement of fit was assessed through likelihood ratio tests comparing models with and without the spillover predictor.

Table 2.1: **Untrimmed raw mean reading times in milliseconds by condition for antecedent, ellipsis and spillover regions, standard errors in parantheses.**

		OVS/amb.	SVO/amb.	OVS/unamb.	SVO/unamb.
<i>A sympathizer ...</i>	np1	1793 (48)	1760 (39)	1830 (41)	1651 (39)
<i>had.sg/pl</i>	aux	519 (17)	474 (8)	499 (12)	474 (10)
<i>the rebels</i>	np2	1021 (28)	976 (28)	913 (23)	921 (27)
<i>according to ...</i>	adj	1041 (26)	1107 (29)	1066 (28)	1135 (31)
<i>decisively supported</i>	vp	892 (23)	887 (24)	868 (22)	900 (26)
...					
<i>substantiate</i>	wh-1	471 (8)	485 (10)	493 (9)	486 (10)
<i>how</i>	wh	423 (7)	427 (7)	422 (6)	434 (7)
<i>so greatly</i>	wh+1	437 (7)	452 (8)	449 (9)	449 (8)
<i>itself</i>	wh+2	578 (15)	564 (15)	591 (16)	584 (18)
<i>the ... commission</i>	wh+3	571 (18)	580 (16)	604 (17)	590 (17)

suggested a reciprocal transformation ($1/RT$). Reciprocal reading times were multiplied by -1000 to make the parameters easier to interpret. Additionally, all data points corresponding to reading times below 150 ms were removed, which resulted in a loss of less than 1% of data in all cases. Effects were judged as significant if $t \geq 2$. Model output is shown in Table 2.2.

Comprehension accuracy

Participants' overall comprehension accuracy was at 90%, though accuracy for experimental items was somewhat lower at 82%. Overall, subjects were most accurate at supplying the wh-pronoun (92% accuracy) and least accurate at verifying statements about the argument structure of the antecedent (72% accuracy), with the rest of the comprehension tests falling in between (86% accuracy). All further analyses were conducted without distinguishing between question types, unless otherwise noted. A linear mixed-effects model was fit to question response times using the same procedure described above for reading times. The analysis revealed no significant effects of the experimental manipulation. An analogous model with reciprocal response time as an additional predictor was fit to response accuracies using a logit link function. The fit showed an effect of response time such that accuracy dropped with increased delay ($\hat{\beta} = -0.13$, $se = 0.03$, $t = -5.18$), as well as a significant word order \times

case marking interaction ($\hat{\beta} = -0.18$, $se = 0.07$, $t = -2.74$), which nested contrasts⁸ revealed to be driven by the OVS/ambiguous condition eliciting more incorrect responses than the SVO/ambiguous condition ($\hat{\beta} = -0.27$, $se = 0.13$, $t = -2.09$). To investigate further, we created a new contrast between questions that queried the role of the arguments in the antecedent and questions that did not. When this distinction was entered into the model,⁹ it turned out to be highly predictive of accuracy ($\hat{\beta} = -0.66$, $se = 0.16$, $t = -4.24$), indicating that questions about argument structure were more difficult to answer than other question types. At the same time, the word order \times case marking interaction was significant ($\hat{\beta} = -0.17$, $se = 0.07$, $t = -2.63$), but there was no three-way interaction. There was thus no indication that comprehension failure for questions targeting argument structure was limited to garden-path sentences. Why answering questions about garden-path sentences should be difficult even when the temporary ambiguity is not targeted remains mysterious for the time being.

Reading times

Table 2.1 shows the mean raw reading times for the analyzed regions of interest. Figure 2.1 shows residual mean reading times for each region of the antecedent. Residualization was carried out by fitting a linear mixed-effects model with region length as a fixed effect and random slopes by subject. Unresidualized reciprocal reading times (see above) were used in the statistical analysis. A main effect of word order appeared at the auxiliary ($\hat{\beta} = 0.03$, $se = 0.01$, $t = 2.07$), such that OVS was processed more slowly than SVO, which is likely due to the additional plural suffix in the OVS conditions. On the second NP, there were main effects of word order ($\hat{\beta} = 0.04$, $se = 0.01$, $t = 3.02$) and case marking ($\hat{\beta} = 0.04$, $se = 0.01$, $t = 3.3$), such that SVO was read faster than OVS and unambiguous sentences were read faster than ambiguous ones. There was also a significant interaction between the factors ($\hat{\beta} = 0.02$, $se = 0.01$, $t = 2.12$), which nested contrasts revealed to be driven by OVS clauses taking longer to read in the presence of ambiguous case marking ($\hat{\beta} = 0.07$, $se = 0.02$, $t = 3.68$). The preverbal adjunct again showed a main effect of word order ($\hat{\beta} = -0.02$, $se = 0.01$, $t = -2.38$); at this position, OVS clauses were read faster than SVO clauses.¹⁰

⁸For this analysis, case marking was treated as nested within word order. One sum contrast compared the two ambiguous conditions, one compared the two unambiguous conditions, and a third one the OVS versus SVO conditions.

⁹The fixed effect of reciprocal response time was removed from this model as it consistently led to convergence failure.

¹⁰Speculatively, this effect may be due to readers trying to make up for lost time after having been slowed down.

Figure 2.1: **Residual reading times for the antecedent regions, extreme values removed.** Error bars represent 95% confidence intervals.

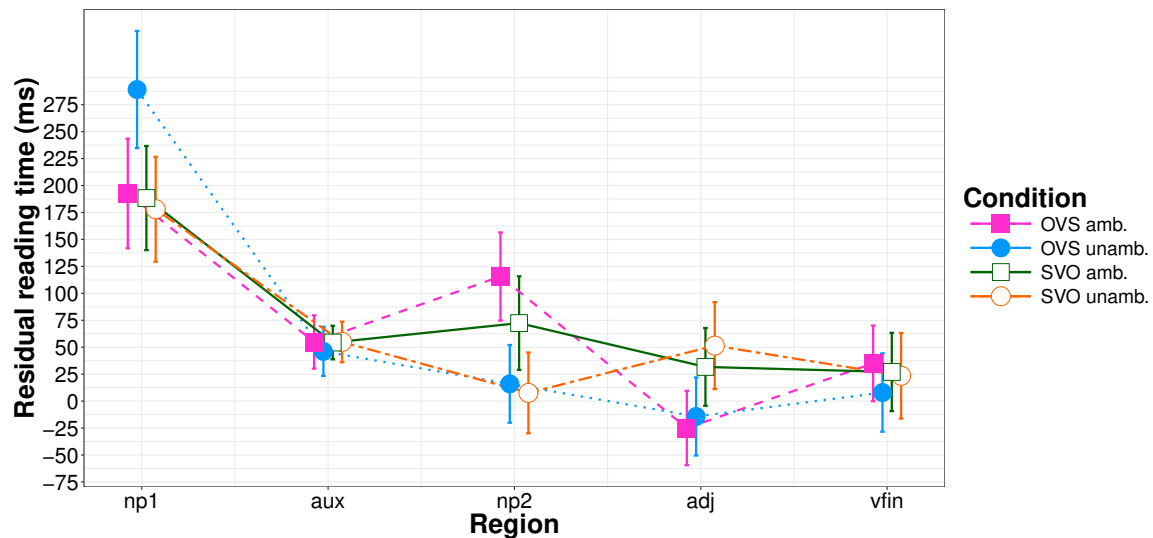


Figure 2.2 shows the mean reading times from the region right before the ellipsis site to three words after the ellipsis site, again in residualized form. No significant effects appeared at the wh-pronoun or in the immediately following region. In the next region (wh+2), there was a main effect of word order ($\hat{\beta} = 0.03$, $se = 0.01$, $t = 2.02$), such that OVS clauses took longer to read than SVO clauses. For this position, closer inspection of the model revealed one very short reading time (177 ms) to be highly influential in the fit, and removing this value resulted in the effect merely approaching significance ($\hat{\beta} = 0.02$, $se = 0.01$, $t = 1.89$). In the third region after the wh-pronoun (wh+3), a word order \times case marking interaction reached significance ($\hat{\beta} = -0.03$, $se = 0.01$, $t = -2.02$), due to the OVS/ambiguous condition being read faster than the OVS/unambiguous condition, with no single condition driving the interaction. During data analysis we noticed that five experimental sentences featured gender-marked pronouns at position wh+2, which presents a possible confound. Adding the presence versus absence of a pronoun as a sum-coded predictor did, however, not change the results found at regions wh+2 and wh+3.

Figure 2.2: Residual reading times for the pre-ellipsis, ellipsis, and spillover regions, extreme values removed. Error bars represent 95% confidence intervals.

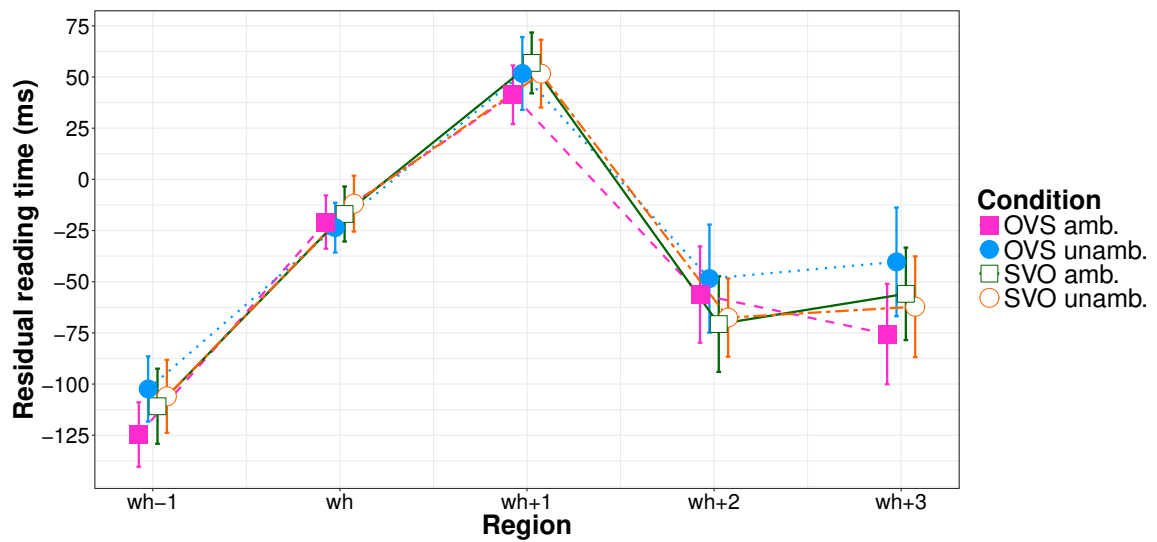


Table 2.2: Coefficient estimates, standard errors and t values for the linear mixed-effects models fit to reciprocal reading times at the indicated regions of interest.

aux			
	Estimate	Std. Error	t value
(Intercept)	-2.19	0.07	-32.82
case marking	0.01	0.01	0.51
word order	0.03	0.01	2.07
spillover	-0.08	0.04	-1.75
case marking:word order	-0.01	0.01	-0.91
np2			
	Estimate	Std. Error	t value
(Intercept)	-1.32	0.08	-17.08
case marking	0.04	0.01	3.30
word order	0.04	0.01	3.02
spillover	0.15	0.02	6.48
case marking:word order	0.02	0.01	2.12
adj			
	Estimate	Std. Error	t value
(Intercept)	-1.17	0.07	-15.96
case marking	0.00	0.01	0.26
word order	-0.02	0.01	-2.38
spillover	-0.01	0.02	-0.83
case marking:word order	0.00	0.01	-0.35
wh+2			
	Estimate	Std. Error	t value
(Intercept)	-1.98	0.10	-20.65
case marking	-0.02	0.01	-1.37
word order	0.03	0.01	2.02
spillover	0.25	0.02	10.22
case marking:word order	0.01	0.01	0.89
wh+3			
	Estimate	Std. Error	t value
(Intercept)	-2.11	0.10	-21.19
case marking	-0.03	0.01	-1.86
word order	-0.00	0.02	-0.12
spillover	0.06	0.02	3.05
case marking:word order	-0.03	0.01	-2.01

One might think that the interaction found at position wh+3 stemmed from occasional processing breakdowns in the OVS/ambiguous sentences. We assume that these would be due to failures in processing the antecedent, which would leave the parser without an adequate retrieval target for the ellipsis. To test this hypothesis, we added the reading time for the second NP, which is expected to reflect the difficulty of the garden path, to the reading time model for position wh+3 on the same trial. While this measure turned out to be a highly significant predictor ($\hat{\beta} = 0.13$, $se = 0.02$, $t = 5.51$), the word order \times case marking interaction also stayed significant and indeed became stronger ($\hat{\beta} = -0.03$, $se = 0.01$, $t = -2.21$). This suggests that while the time spent processing the garden-path influences retrieval difficulty, there are factors above and beyond this measure which determine processing effort at the ellipsis site. In a further test, we added reading times for both the second NP and position wh+3 to the response accuracy model reported above. The reasoning behind this was that processing failure at either position could lead to incorrect responses. Adding these parameters did, however, not change the result. We also compared the median reading time in the OVS/ambiguous condition for position wh+3 with the overall median reading time for the experimental items. The difference lay within reasonable bounds (439 ms, $se = 18$ ms vs. 473 ms, $se = 2$ ms), indicating that very short RTs from processing failures were not pushing down the median. Congruently with this, a visual inspection of a density plot of RTs at position wh+3 did not indicate a mode or tail of fast reading times, nor did Hartigan's Dip Test (Hartigan & Hartigan, 1985) yield any evidence for bimodality. Finally, we removed all trials with incorrect responses to the comprehension test, which amounted to 18% of the data for position wh+3, and refit our model. Note that an incorrect answer does not necessarily mean that parsing failed; misinterpretations could, for instance, arise from fragments of discarded analyses in memory (see below). Nevertheless, the results of the comprehension test are the only pertinent measure available to us. With one fifth of data removed, the word order \times case marking interaction stayed near the significance threshold ($\hat{\beta} = -0.02$, $se = 0.01$, $t = -1.62$) and became marginally significant when antecedent reading time was added as a predictor ($\hat{\beta} = -0.03$, $se = 0.01$, $t = -1.86$). The loss of significance is not particularly unexpected given the loss of statistical power incurred by removing data. To our minds, these results do not indicate that processing failure was a factor in decreasing reading times for the OVS/ambiguous condition.

Discussion

The expected garden-path effect for the antecedent appeared one region later than predicted, at the second NP, showing that the experimental manipulation was successful. While no effects were found at the ellipsis site itself, OVS antecedents led to longer reading times two

regions downstream from the wh-pronoun. Note that this cannot be explained by a ‘global spillover effect’ from the antecedent: earlier regions did not show the pattern, and there is no reason to assume that antecedents in the OVS/unambiguous condition were extremely difficult to process. Furthermore, an interaction between the experimental factors appeared at position wh+3, albeit in a surprising form: sentences in the OVS/ambiguous condition were read faster than those in the OVS/unambiguous condition, with the two SVO conditions lying in between. We assume that the observed pattern reflects delayed processing of the ellipsis, either as the consequence of subjects ‘tapping’ the space bar at fixed time intervals (Witzel, Witzel, & Forster, 2012; see discussion below) or as spillover that was not factored out by the statistical model. As the OVS/ambiguous condition was responsible for the garden-path effect within the antecedent clause, the processing advantage is unexpected with regard to the reconstruction hypothesis, which had predicted the same pattern to reappear at the ellipsis site. The result is also not straightforwardly explained by a pointer-based approach, which would have predicted no differences between the conditions. We will argue below that what we are observing at positions wh+2 and wh+3 is the interaction of two factors: antecedent-ellipsis mismatch and memory trace reactivation through reanalysis.

German word order and antecedent-ellipsis mismatch

As we’ve pointed out in the introduction, German subordinate clauses are required to be verb-final¹¹ while main clauses invariably have the finite verb in second position. As the sluicing structures in the present study appeared in subordinate clauses, all antecedent clauses would therefore have had to be verb-final instead of verb-second to be compatible with the gap. Given that sluicing is still perfectly acceptable in all of our stimuli, we seem to be seeing a case of ‘acceptable ungrammaticality’ (Frazier, 2008). Both SVO and OVS antecedents were, to use the terminology of Arregui, Clifton, Frazier, and Moulton (2006), ‘flawed’, but possibly not in the same way.

OVS order in German main clauses can be derived through topicalization, with the object occupying the so-called *Vorfeld* (‘prefield’, e.g., S. Müller, 2005).¹² As this strategy is not available in subordinate clauses, non-canonical word orders must be derived via *scrambling*, which moves constituents within the so-called *Mittelfeld* (‘middle field’, e.g., Hinterhölzl,

¹¹The only exception to this rule occurs when the verb takes a sentential complement, which was not the case in our experiment.

¹²The *Feldertheorie* of German sentence structure was first developed by Drach (1937), and is also known as the *Topological Model*.

2006). The slightly simplified examples in (10) illustrate this. SOV order in (10a) is unproblematic, but scrambled OSV in (10b) is, at the very least, highly marked.¹³

- (10) Die Regierung konnte nicht nachweisen, ...
 the government could not substantiate
- a. **SOV subordinate clause**
 ... wie die Rebellen einen Sympathisanten unterstützt hatten.
 how the rebels a sympathizer.acc supported had.pl
- b. **OSV subordinate clause**
 ... ?? wie einen Sympathisanten_i die Rebellen t_i unterstützt hatten.
 how a sympathizer.acc the rebels supported had.pl

The Recycling Hypothesis proposed by Arregui et al. (2006) predicts that ellipses are more difficult to process the more the antecedent mismatches the ellipsis site. Arregui et al. assume ‘repair’ operations as the source of the difficulty. Assuming the verb-second antecedents have already been partly repaired by moving the verb to the end, (10b) would still need to be transformed into an SOV structure like (10a), presumably by reversing the movement operation. The increased reading times for sentences with object-initial antecedents observed at position wh+2 would be expected under the assumption that the mismatch between an OVS antecedent and an SOV sluice is greater than for SVO antecedents, where the repair process does not need to change the order of the arguments.

Two alternative suggestions made by an anonymous reviewer merit discussion. One is that the processor simply fills the ellipsis site with a verb-second clause, deriving a structure that would have no grammatical surface equivalent. There would be no reason to invoke the Recycling Hypothesis in this case, and the OVS disadvantage would need to be explained either by constraints on topicalization or possibly by invoking working memory factors. Both of these possibilities present problems. It has been found that surprising or unusual stimuli lead to better recall performance (Hirshman, Whelley, & Palij, 1989), which would lead us to expect that the more uncommon OVS antecedents should be easier instead of more difficult to retrieve. Additionally, the claim that ungrammatical structures can be derived during ellipsis processing seems extreme given that the observed effects can be explained through other means. The reviewer’s second suggestion is that garden-pathing in the antecedent might result in its memory representation being more difficult to access, allowing a slower discourse-based mechanism like Murphy’s (1985) to dominate during

¹³Apart from not being licensed by information structure, moving the object in (10b) also violates a constraint dictating that definite noun phrases should appear before indefinite ones (see G. Müller, 1999 for an optimality-theoretic account).

processing. However, seeing that *unambiguous* OVS antecedents also led to longer reading times at position wh+2, this does not seem like a plausible alternative to us.

Antecedent reactivation through reanalysis

A reviewer points out that there is some evidence that initial misinterpretations of garden-path sentences persist beyond the point of disambiguation, leading to structural priming (van Gompel, Pickering, Pearson, & Jacob, 2006), systematic errors during paraphrasing (Patson, Darowski, Moon, & Ferreira, 2009) and in comprehension tests (Christianson, Hollingworth, Halliwell, & Ferreira, 2001), as well as competition effects when late-arriving plausibility information contradicts the initial parse (Slattery, Sturt, Christianson, Yoshida, & Ferreira, 2013). One explanation for these effects is that the initial parse of the sentence remains active in memory to some degree even after it has been discarded. In the case of our experiment, if a remnant of the discontinued subject-initial analysis remains behind in the OVS/ambiguous condition, it might be conceivable that this memory trace is considered as a possible antecedent for the ellipsis, possibly blocking access to the ‘real’, reanalyzed antecedent. Research on agreement processing, reflexives and subject-verb dependencies has shown that such memory interference may turn out to make processing easier or more difficult, depending on the phenomenon under study and the exact setup of the experiment (see Jäger, Engelmann, & Vasishth, 2017 for a review). While the observed speedup in the current study may, in principle, be explained through facilitative interference, the results of Martin and McElree (2009) suggest that the availability of multiple candidate antecedents does not influence the time-course of ellipsis processing in any way. As it is unclear why the interference effect should be visible in our experiment but not in theirs, we will present an alternative explanation of our results.

We suggest that the pattern at position wh+3 should be analyzed in terms of a reactivation of the antecedent’s memory trace that outweighs the mismatch penalty created by the word order manipulation. As explained in the introduction section, the cue-based retrieval parser of Lewis and Vasishth (2005) incorporates the assumption that syntactic phrases are stored in working memory as chunks. If a chunk is retrieved in order to make an attachment, its activation level increases, which makes subsequent retrievals easier. A reanalysis such as the one required for sentences in the OVS/ambiguous condition should reactivate the antecedent’s memory chunk as its structure needs to be changed. Later, at the ellipsis site, it should thus be retrieved faster than the other types of antecedents, to which reanalysis has not applied.¹⁴

¹⁴This presupposes that trace decay has not reduced the activation of the antecedent to zero in any case by the time the ellipsis is processed. The model of Lewis and Vasishth (2005) assumes that the activation of chunks that have been reaccessed is higher even after complete decay.

The mismatch effect explained above can also be accounted for through an extension of the Lewis and Vasishth (2005) model: If the *wh*-pronoun sets retrieval cues for a verb-final antecedent in order to match the local clausal configuration, there will be no matching chunk in memory. In order to be able to complete the retrieval, the processor may then attempt to retrieve chunks which do not match the cues perfectly, such as the main clauses in the current study. Due to the matching relative order of subject and object, an SVO chunk may resonate more strongly with the SOV cue than one with OVS word order, as schematized in (11).

- (11) a. **OVS antecedent, resonates weakly with SOV cue (O-S \neq SO)**
 [Einen Sympathisanten hatten die Rebellen unterstützt],_{OVS} ...
 A sympathizer had.pl the rebels supported
- b. **SVO antecedent, resonates more strongly with SOV cue (S-O \sim SO)**
 [Ein Sympathisant hatte die Rebellen unterstützt],_{SVO} ...
 A sympathizer had.pl the rebels supported
wie in subordinate clause sets SOV cue
 ... aber die Regierung konnte nicht nachweisen, wie []_{SOV} ...
 but the government could not substantiate how

A lower retrieval latency would then be expected for SVO chunks, thereby predicting the observed OVS disadvantage at position *wh*+2.¹⁵ The reactivation/mismatch approach is thus able to account for the observed pattern of results, but due its status as a *post-hoc* argument is in need of further empirical validation.

One might think of yet another explanation for the result, namely that reconstruction *is* taking place and that syntactic priming is responsible for the advantage in the OVS/ambiguous condition. However, such an approach would not fit with the fact that the antecedent's structure is, strictly speaking, incompatible with the word order required at the gap: As the derivations of main and subordinate clauses involve different steps, it is not obvious what exactly would be primed. One would have to make a very specific set of assumptions: First, the parser would need to blindly reconstruct the syntax of the antecedent at the ellipsis site *before* checking for possible mismatches, similarly to the anonymous reviewer's suggestion that was discussed earlier. Secondly, garden-path sentences would need to prime their final structure more strongly than unambiguous controls, which to our knowledge has not been demonstrated to date. Ambiguous/OVS antecedents would then initially gain an advantage through increased priming while both kinds of OVS antecedents would be disadvantaged during the mismatch checking phase.

¹⁵In order to derive grammatical structures, repair processes that change the word order to verb-final would still need to apply after retrieval.

Sluicing and predictive processing

We believe that one additional result is worth mentioning, even though it was only arrived at *post hoc*. It fits with the proposal by Yoshida et al. (2013) that predictive processing may be involved in the interpretation of sluicing structures. Yoshida et al. compared sentences in which it was either possible or impossible to analyze a specific wh-phrase as part of a sluice. The evidence suggested that as soon as the wh-phrase in question was encountered, the parser started building a sluicing structure, presumably because it is preferred over other possible continuations.

We took the implication of predictive processing as an incentive to analyze reading times for the region directly *preceding* the wh-pronoun in our own experiment: If sluicing is the preferred continuation after a wh-pronoun has been encountered, it is not unlikely that it will also rank fairly highly before that point. This is especially likely given that subordinate clauses in German require a comma, which was thus present in the pre-wh region in all of our stimuli, excluding a vast range of alternative continuations that would have been likely in Yoshida et al.'s materials.

The fitting of a linear mixed-effects model (see above) at position wh-1 revealed a significant interaction between word order and case marking ($\hat{\beta} = -0.03$, $se = 0.01$, $t = -2.3$) which had the same sign as the one observed at position wh+3.¹⁶ Table 2.3 shows the model output. However, unlike at the later position, nested contrasts showed that the interaction was driven by the OVS/unambiguous condition being read more slowly than the SVO/unambiguous condition ($\hat{\beta} = 0.04$, $se = 0.02$, $t = 2.24$), even though the numerical pattern in raw reading times was the same as for position wh+3. We have no ready explanation for this finding. Speculatively, a heuristic may be used to estimate the fit between the sluice and the antecedent. Such a heuristic might work better when case is overtly marked, and might operate more quickly when word order is canonical. In our opinion this kind of predictive strategy makes it unlikely that processing proceeds according to the priming-based account described above, in which local constraints do not influence the initial structure assignment for the ellipsis.

To further investigate the notion that a sluice was the expected structure in our materials, we ran a sentence completion study with 35 new participants. It has been suggested that the speech production system may be responsible for generating linguistic expectations in comprehension (Pickering & Garrod, 2007). As sentence continuation preferences have been shown to be predictive of processing difficulty in self-paced reading (Smith & Levy, 2011), we assume that a preference for sluicing continuations in our reading study should translate into

¹⁶As a sanity check, we also analyzed reading times at position wh-2, finding no significant effects.

a corresponding preference in sentence completions. The stimuli consisted of the 32 sentences used in the current reading study, along with 32 sentences from a different experiment and 96 fillers. Sentences were presented using a modified version of Linger’s masked auto-paced reading (otherwise known as rapid serial visual presentation or RSVP). The stimuli from the current study were cut off right before the ellipsis site and participants were asked to complete the sentences using the first continuation that came to mind. Due to the nature of the presentation, participants could not reread the sentences while they were typing their continuation. Results showed a total of only 5% sluicing continuations. Another 45% of continuations were non-sluiced *wh*-clauses, followed by *if*-clauses at 17% and *that*-clauses at 7%. Assuming that this pattern is not due to idiosyncrasies of the production system, the observed outcome casts some doubt on the assumption that a sluicing continuation was, in fact, highly expected in our stimuli. However, subjects in the production experiment could choose their preferred continuation freely, which may conceivably have led to more conscious deliberation on their part. It is entirely possible that sluicing is only one of several possible continuations which are preactivated during reading, which might be enough to explain the findings of Yoshida et al. (2013) and the interaction we observed at position *wh*–1 in the self-paced reading study. Despite the limited scope of the production experiment, given the earlier findings by (Smith & Levy, 2011), we feel that it was important to investigate whether the predictive processing seen in comprehension maps directly onto language users’ preferences in production. This is apparently not the case under the conditions tested here.

Table 2.3: **Coefficient estimates, standard errors and t values for the linear mixed-effects model fit to reciprocal reading times at region *wh*–1.**

<i>wh</i> –1			
	Estimate	Std. Error	t value
(Intercept)	–2.10	0.06	–35.17
case marking	–0.02	0.01	–1.60
word order	0.01	0.01	0.93
spillover	0.36	0.02	15.21
case marking:word order	–0.03	0.01	–2.30

General Discussion

The current experiment investigated the processing of a sluicing construction in cases where the antecedent is a garden-path structure, in this instance a clause with a subject/object ambiguity. We observed reduced reading times for sentences with garden-path antecedents three regions downstream from the ellipsis as well as directly before the ellipsis. Furthermore, there was an overall pattern of elevated reading times in the spillover region for antecedents that mismatched the canonical word order of the ellipsis site. Our results are best compatible with accounts of ellipsis resolution that can be implemented in the form of a memory pointer mechanism (Frazier & Clifton, 2001, 2005; Martin & McElree, 2008), which would need to be augmented to account for reactivation assumed by the cue-based retrieval parser of Lewis and Vasishth (2005). The evidence for a mismatch effect is in line with the predictions of the Recycling Hypothesis proposed by Arregui et al. (2006). However, given that we have observed no evidence for reconstruction in our experiment, we do not subscribe to Arregui et al.’s assumption that ‘flawed’ antecedents are ‘repaired’ in a way that is similar to syntactic reanalysis (p. 242). The mismatch effect may be better approached along the lines of the wh-pronoun setting a retrieval cue for an antecedent that matches the word order requirements of the local clause, opting for the closest candidate upon failure. Alternatively, one could follow the proposal of Kim et al. (2011), in which ellipses with non-canonical antecedents violate parsing heuristics that are based on construction frequency and expectation. Under an approach without reconstruction, we would claim that it is not a parsing heuristic that is violated, but a local expectation as to what an antecedent targeted by retrieval should look like. If the expectation were global, no mismatch effect would be expected, given that the antecedent has already been encountered in the input. The local expectation account fits with the pattern observed by Yoshida et al. (2013) as well as with the effect found in the pre-sluice region (wh-1) in the current study.

Still, why did we observe a pattern in which the experimental manipulation seemed to have an effect before and after, but not *at* the ellipsis site? We assume that this is due to either insufficient statistical power, to our subjects’ reading strategies, or both. Power is always an issue when effect sizes are as small as in the current study: the mean reading time difference between the unambiguous/OVS and the ambiguous/OVS conditions at position wh+3 was only 30 ms. Given this value and the associated standard errors, the post-hoc power to detect a real effect was at 45%, which is comparable to Frazier and Clifton’s (2000) study, where the computation yields 43% post-hoc power.¹⁷ The bottom line is that sample size needs to

¹⁷Note that this is not the *true* power of the experiments, which depends on the unknown true effect size.

be significantly increased in order to convincingly argue that there really is *no* effect of the manipulation, even though this might be construed as trying to ‘force significance’.

The concern related to reading strategies comes from the fact that while non-cumulative self-paced reading more closely resembles data from natural reading than the cumulative variant does (Just et al., 1982), it is by no means certain that subjects will not adopt a ‘wait and see’ strategy at least on some trials, meaning that they will press the button at a fixed rate and only then start processing. Witzel et al. (2012), suspecting such rhythmic ‘tapping’ in their data, tried to remove its influence by calculating the standard deviation of the response time by subject and excluding the participants with the smallest variability, which did, however, not change their statistical result. The authors conclude that either ‘tapping’ was not a factor in their data or their method was not suitable to account for it, leaving the issue for future research. We will do the same here.

There is also a slightly different explanation for the delay we observed, namely that subjects *did* process the words the words as they were revealed, but postponed the processing of the ellipsis until they had more information. Such a strategy might make sense considering that an embedded question (i.e., an interrogative clause that serves as a complement, as in . . . , *but the government could not substantiate how, . . .*) in itself usually imparts no relevant information apart from the fact that some piece of information is missing. As the contents of the spillover region put this information in context (. . . , *because/so that/even though/until . . .*), the relevance may have become apparent, causing the observed processing pattern.

A final objection to our study would be that there was no control condition without ellipsis. It should be noted that it is extremely difficult to create closely matched controls for our sentences, given that possible continuations are limited to complement clauses, which usually feature more than one word. Other studies on ellipsis processing also lack controls (e.g., Frazier & Clifton, 2000, 2005 [except Experiments 2 and 3], Poirier, Wolfinger, Spellman, & Shapiro, 2010), leaving open the possibility that any observed effects do not actually stem from the antecedent being recovered due to a perceived gap in the sentence but from some other mechanism. While this criticism can be met by pointing to the localization of the effects, as well as to the unavailability of a plausible alternative explanation, it would be desirable to include controls in future studies to strengthen the conclusions drawn from the data.

Further investigations into the interaction between antecedent ambiguity and ellipsis processing are already underway in our laboratory. We are currently aiming to find further evidence for the reactivation effect using different kinds of temporary ambiguities and ellipses, as well as experimental procedures other than self-paced reading (e.g., eye tracking).

Acknowledgments

To author wishes to thank Shravan Vasishth, Lena A. Jäger, Barbara Hemforth, the Vasishth Lab team and the audience at CUNY 2015 for helpful comments and suggestions, as well as Johanna Thieke for assistance with data collection.

CHAPTER 3

DOES ANTECEDENT COMPLEXITY AFFECT ELLIPSIS PROCESSING? AN EMPIRICAL INVESTIGATION

Dario Paape, Bruno Nicenboim, and Shravan Vasishth

GLOSSA: A JOURNAL OF GENERAL LINGUISTICS, VOLUME 2, ISSUE 1

Abstract

In two self-paced reading experiments, we investigated the effect of changes in antecedent complexity on processing times for ellipsis. Pointer- or ‘sharing’-based approaches to ellipsis processing (Frazier & Clifton, 2001, 2005; Martin & McElree, 2008) predict no effect of antecedent complexity on reading times at the ellipsis site while other accounts predict increased antecedent complexity to either slow down processing (Murphy, 1985) or to speed it up (Hofmeister, 2011). Experiment 1 manipulated antecedent complexity and elision, yielding evidence against a speedup at the ellipsis site and in favor of a null effect. In order to investigate possible superficial processing on part of participants, Experiment 2 manipulated the amount of attention required to correctly respond to end-of-sentence comprehension probes, yielding evidence against a complexity-induced slowdown at the ellipsis site. Overall, our results are compatible with pointer-based approaches while casting doubt on the notion that changes in antecedent complexity lead to measurable differences in ellipsis processing speed.

Introduction

Murphy (1985) observed elevated whole-sentence reading times for the second clause in (12b) as compared to (12a).

- (12) a. Jimmy swept the floor. Later, his uncle did too.
b. Jimmy swept the tile floor behind the chairs free of hair and cigarettes. Later, his uncle did too.

In these examples, the sentence *Later, his uncle did too* contains a verb-phrase ellipsis, such that the auxiliary *did* is taken to carry the meaning of the entire verb phrase of the preceding clause. Murphy explains his experimental findings by assuming a process that copies the antecedent string into the ellipsis site. The assumption that it should take more time to transfer larger amounts of information is rather straightforward if one assumes a constant rate of throughput. Since the copied antecedent meaning is more complex in (12b) than in (12a), it is not surprising that processing time for the ellipsis should increase, given that the predication made of Jimmy's uncle becomes more complex as well.

Clearly, ellipsis is an anaphoric device, and thus superficially similar to pronouns like *he* or *she*. It can thus be assumed that some sort of memory retrieval is initiated when the ellipsis site is encountered. However, Murphy's invocation of a copying process implies that information is duplicated, unlike in the case of pronouns, which simply refer back to an existing discourse entity.¹ Indeed, the uncle's sweeping in (12) is not identical with Jimmy's own sweeping, but refers to an independent event taking place at a different point in time.

Later studies have found no antecedent complexity effects on ellipsis processing. Using self-paced reading and a speed-accuracy trade-off (SAT) procedure, respectively, both Frazier and Clifton (2000) and Martin and McElree (2008) failed to find any evidence of longer antecedents leading to slowed processing at an ellipsis site. Based on their earlier findings, Frazier and Clifton (2001) conclude that copying is 'cost-free', that is, it involves no measurable computational effort. Martin and McElree (2008) propose to do away with the copying metaphor and instead think of ellipsis as a pointer into memory. The reasoning behind the latter approach is that it is sufficient to create a link between an existing representation of the antecedent and

¹Note that the idea that pronouns are, underlyingly, copies of their antecedents, used to be fairly common in theoretical linguistics, but has long since become marginalized (Hankamer & Sag, 1976). Pronouns cannot simply be replaced by their antecedent noun phrases in most contexts, as shown by sentences like *The caterpillar_i will be able to fly when it_i/the caterpillar_i is a butterfly* (Grinder & Postal, 1971, p. 269).

the ellipsis site, much like creating a shortcut to a computer file, rather than creating a duplicate. This view is equivalent to what Frazier and Clifton (2005) call ‘structure sharing’: in essence, one and the same phrase is attached in two places at once. Under this view, ellipsis is no different from pronouns in that it simply refers back to an existing linguistic entity.

In fact, having failed to find an antecedent complexity effect in a second experiment where a sentence intervened between ellipsis and antecedent, Murphy (1985) also introduces the concept of a memory pointer. He argues that comprehenders have both a structure-based and discourse-based mechanism for recovering an antecedent at their disposal. The latter is conceived of as a memory pointer and thus not subject to complexity effects (Murphy, 1985, p. 293) while the former is argued to involve word-by-word copying. Having a clause intervene between ellipsis and antecedent arguably forces readers to fall back on the discourse-based pointer mechanism, presumably because increased distance makes the syntactic or semantic representation of the antecedent unrecoverable.

The pointer/sharing approach runs contrary to the point made above about the independence of ellipsis and antecedent. It involves, to use a term from programming, a ‘shallow copy’ of the antecedent: the ellipsis site is interpreted by looking up the stored value from memory, but does not contain any information besides the pointer. A ‘shallow copy’ is also used in Frazier and Clifton’s (2001) account, thus rendering it equivalent in terms of predictions to the pointer/sharing view. Murphy (1985), on the other hand, assumes a ‘deep copy’, where the information is present in both positions, as the basis for interpretation for the ellipsis. This latter conception is also often implicitly assumed in theoretical linguistics, especially if the ellipsis site is assumed to contain syntactic information (e.g., Merchant, 2001; Williams, 1977)². As Martin and McElree (2008, p. 882) explicitly assume that the antecedent’s memory representation is accessed based on its “required morpho-syntactic, semantic, referential, and pragmatic properties”, we will not subscribe to or compare any accounts which claim that ellipsis processing is exclusively syntax-, semantics- or discourse-based. In fact, the question is orthogonal to the issue of antecedent complexity, as an increase in complexity on any of the aforementioned levels will usually be accompanied by increased complexity on the other levels. However, it does strike us as being most likely that the sentence processor makes use of as much information as it can, irrespective of the source, in order to successfully complete the retrieval.³

²See e.g., Dalrymple, Shieber, and Pereira (1991) for examples of empirical problems caused by a purely syntactic view.

³An anonymous reviewer notes that a recent study by Cai et al. (2013) showed no evidence that ellipsis leads to structural priming, which can be taken to imply that syntactic information is not relevant. However, Xiang et al. (2014) did find a structural priming effect, and indeed the mere retrieval of a syntactic representation of the antecedent does not necessarily imply that any priming should be expected, given that – arguably – no structure is actively computed.

Both Murphy's (1985) experiment and the studies of Frazier and Clifton (2000) and Martin and McElree (2008) have been criticized by other scholars. Tanenhaus and Carlson (1990) note that Murphy's long antecedents may have contained temporary attachment ambiguities, while Phillips and Parker (2014) point out methodological flaws in the two more recent studies (see also Paape, 2016, p. 3), which are discussed below.

Frazier and Clifton (2000) tested sentences like (13) in a self-paced reading paradigm.

- (13) a. Sarah left her boyfriend last May. Tina did too.
b. Sarah got up the courage to leave her boyfriend last May. Tina did too.

There were only twelve items in this study, six of which were accompanied by comprehension questions that did not target the interpretation of the elliptical clause. In addition to the sentences shown above, there were two additional versions of the complex – that is, (13b) – variant of each item in which the two clauses were connected by the conjunction *and*, which means that each subject tested contributed four data points for every cell of the design. No significant effects of antecedent complexity on reading times for the elliptical clause were found, but there was a trend of 50 ms (SE: 28 ms) towards the segment *Tina did too* being read more slowly when the antecedent was complex.

As pointed out by (Phillips & Parker, 2014, p. 91), this result raises at least three major concerns. First and foremost, even though sixty subjects participated in the experiment, there could be a loss of power due to the relatively small number of observations from each participant. Moreover, and this applies to the studies of Murphy (1985) and Martin and McElree (2008) as well, measuring at the end of a sentence may introduce confounds from so-called wrap-up effects (Just & Carpenter, 1980). The basic observation is that readers generally spend more time reading sentence-final regions, as well as triggering more and longer saccades in eye-tracking, which has been attributed to an overhead of comprehension processes that were not carried out during the reading of the stimulus. Whatever the exact nature of these processes, it is conceivable that their application may mask an effect of antecedent complexity on ellipsis processing. A related criticism is connected to spillover effects. Especially in self-paced reading, the effect of an experimental manipulation often only appears one or two regions downstream from where it would be expected, indicating that subjects do not finish processing each presentation region before continuing to the next one. It is thus possible that readers were still busy integrating the antecedent into the first clause when they encountered the second one, and that any observed effect of complexity is due to processing spillover.

The final concern is about the effect of task demands on reader behavior. Studies have repeatedly shown that readers adapt to experimental demands: they may fail to carry

out processing steps necessary for reference assignment (Foertsch & Gernsbacher, 1994), underspecify syntactic attachments (Swets et al., 2008) and leave quantifier scope ambiguities unresolved (Dwivedi, 2013) in the absence of explicit motivation in the form of well-designed comprehension tests. As reported above, Frazier and Clifton (2000) did not query which meaning their readers had derived for the elliptical clause, and in fact did not ask any comprehension questions in half of the experimental trials. Martin and McElree (2008) used a speed-accuracy trade-off paradigm which involved end-of-sentence grammaticality judgments. These could, however, be made correctly by simply monitoring the animacy of the unelided subject of the VP ellipsis, a strategy which does not require any deep processing of the elided part of the clause (cf. Phillips & Parker, 2014, p. 91).

Given the concerns raised above, we feel that the issue of antecedent complexity effects in ellipsis processing has not yet been satisfactorily resolved. In our own studies reported below, we attempt to address the problems noted by the aforementioned critics. Experiment 1 resolves the issue of end-of-sentence measurements, in addition to using a non-elliptical control condition, while Experiment 2 directly tests for an influence of task demands on antecedent complexity effects. First, however, yet another perspective on the possible effects of antecedent complexity on ellipsis processing will be introduced; it predicts that instead of slowing down the interpretation process, increased complexity of the antecedent should result in a speedup.

Hofmeister (2011) investigated the processing of cleft sentences, which contain a filler-gap dependency between the clefted constituent and the position it was extracted from. In (14), the phrase *a (...) communist* is the object of the verb *banned*, and thus has to be retrieved from memory when the verb is read to compute the meaning of the clause.

- (14) It was [a communist]/[an alleged communist]/[an alleged Venezuelan communist] who the members of the club banned from ever entering the premises.

In a self-paced reading study, Hofmeister found that reading times right after the verb *banned* decreased with the complexity of the filler phrase. Further experiments showed that increasing the semantic specificity of the antecedent also decreased processing times at the gap when string length was kept constant (*which person* vs. *which soldier*), but that making the filler difficult to process (*the lovable military dictator*) resulted in a slowdown rather than a speedup. Hofmeister concludes that more elaborate descriptions of retrieval targets aid memory access as long as they are ‘typical’ (*ruthless military dictator* showed an advantage over *dictator*). He proposes that features which are closely associated (*ruthless – dictator*; *wealthy – celebrity*) will speed up access to the memory target because activation spreads from feature to feature.

Coming back to ellipsis, if the event description encoded by the retrieval target – that is, the verb phrase in (12) – becomes more elaborate, it should become easier to access. Informally, when a reader of (12b) encounters the word *did*, remembering that Jimmy did something involving hair and cigarettes might facilitate access to the sweeping event described by the antecedent. While this is precisely the opposite of what Murphy (1985) observed, it is possible that any advantage due to elaboration was lost due to idiosyncrasies of the items used in his study. In Murphy’s (12a), it does not matter whether the floor was still dirty when Jimmy’s uncle swept it, while in (12b) it clearly was, which requires a laborious inference on part of the reader.

In the two self-paced reading studies we present in this paper, we investigated the processing of ellipses with antecedents of varying complexity. In order to broaden the scope of the inquiry, Experiment 1 focused on German instead of English. Since VP ellipsis does not exist in German, stimulus sentences in this experiment contained a construction known as bare argument ellipsis, also called ‘stripping’. Experiment 1 improved upon previous studies that did not feature control conditions without ellipsis – Martin and McElree (2008) being a notable exception – and featured a subset of comprehension questions that directly targeted the interpretation of the elliptical clause. Experiment 2 addressed the concern originally raised by Phillips and Parker (2014) that superficial processing may have played a role in the studies of Frazier and Clifton (2000) and Martin and McElree (2008). To this aim, we manipulated the types of comprehension questions that participants had to answer, much like Swets et al. (2008) did when investigating the resolution of temporary syntactic ambiguity.

Experiment 1

Bare argument ellipsis or ‘stripping’ deletes an entire clause, with the exception of one constituent, plus an adverb in some cases (Hankamer & Sag, 1976). A German example is given in (15), where the second of the conjoined clauses is understood to mean *John wanted to jump over the fence as well*.

- (15) Peter wollte über den Zaun springen und Johann ebenfalls.
Peter wanted over the fence jump and John as.well
‘Peter wanted to jump over the fence and John (did) as well.’

Stripping targets constituents which are larger than VP, as evidenced by the fact that the modal is deleted along with the lexical verb. Apart from this, we know of no reason why the processing of stripping constructions should differ fundamentally from that of VP ellipsis

in English, other than that cues for a different kind of retrieval target are set, and that the cuing element in this case is an adverb rather than an auxiliary. As with VP ellipsis, when the gap site is identified, the processor needs to look for a suitable antecedent whose meaning (or structure) the gap is to be identified with. In this example, the antecedent consists of the string *wollte über den Zaun springen*, ‘wanted to jump over the fence’.

Materials

A sample stimulus from Experiment 1 is shown in (16). Diamonds indicate the boundaries of presentation regions during the experiment. The study employed a 2×2 design with the experimental factors antecedent complexity (simple vs. complex) and elision (ellipsis vs. control). A total of 28 items were created. The stimuli are listed in Appendix II. Ninety filler items featuring a variety of constructions were also presented during each experimental session.

(16) **Simple antecedent**

Die Armee ◊ räumte ◊ einige wichtige Feldlager ◊ ...
the army cleared some important field camps
‘The army cleared some important field camps ...’

Complex antecedent

Die
the
Armee ◊ ließ ◊ nach dem Gefecht ◊ einige wichtige Feldlager ◊ räumen ◊ ...
army let after the battle some important field camps clear
‘The army had some important field camps cleared after the battle ...’

Continuation

... und ◊ der kluge Befehlshaber ◊ der Aufständischen ◊ ...
and the clever commander of the insurgents
‘... and the clever commander of the insurgents ...’

Ellipsis

... ebenfalls ◊ ...
as.well

Control

... rückte vor ◊ ...
advanced

Continuation

... ohne
without
◊ allerdings ◊ die Sicherung ◊ seiner Versorgungslinien ◊ zu vernachlässigen.
however the protection of his supply lines to neglect
'... without, however, neglecting the protection of his supply lines.'

All experimental sentences featured the same structure, namely an antecedent clause connected to another clause via the conjunction *und*, 'and'. The critical region is the final word of the second clause, which is either the adverb *ebenfalls*, 'as well', or an intransitive lexical verb. The adverb signaling the ellipsis remained the same across all items while the verbs in the control conditions differed. Antecedent complexity was manipulated by adding a modal verb or auxiliary and an adjunct to the simple version of the first clause.⁴ The sentence continues after the critical region in order to prevent wrap-up effects due to periods and allow for spillover.

Participants

Sixty native speakers of German participated in the experiment. These were recruited from the Vasishth Lab's subject pool at the University of Potsdam, which is administrated and maintained through ORSEE (Greiner, 2015). Each subject was either paid 6 € or received course credit. Informed consent from the participant was obtained before each experimental session. The experiment complied with the June 1964 Declaration of Helsinki (carried out by the World Medical Association and entitled "Ethical Principles for Medical Research Involving Human Subjects"), as last revised. In accordance with German NSF (DFG) guidelines, for experiments with unimpaired adult populations, the ethics approval is required by the Principal Investigator (in this case, Prof. Dr. Shravan Vasishth).

⁴The causative verb *lassen*, 'let', is, strictly speaking, neither a modal verb nor an auxiliary, even though it is known to behave like one of the former in most respects (Bader, 2014). For the purposes of the present study, the central requirement was that *lassen* embeds an infinitive and is deleted along with it in ellipsis. Other items featured more typical modals such as *wollen*, 'want', *müssen*, 'must', and *sollen*, 'should', or the auxiliary *haben*, 'have'.

Procedure

The experimental stimuli were presented in a latin-square design using the **Linger** software written by Douglas Rohde (Rohde, 2003), along with the filler items. Presentation order was randomized at runtime. Participants were instructed to read silently at their normal pace. Each trial started with a white screen that was displayed for 1000 ms and that could not be skipped. The sentence was then shown in masked form, that is, with all characters except spaces replaced by underscores (`_`). Participants pressed the space bar to replace the underscores with the corresponding regions of the sentence, displayed in 20 pt Courier New font. Presentation was non-cumulative, that is, previous regions reverted back to underscores upon continuation. Times between button presses were recorded. After every sentence, a statement was shown that participants were required to judge as being either true or false, based exclusively on the information given by the stimulus. For instance, a subject reading the simple/ellipsis version of (16) would have been required to judge the statement *A clever commander had some important field camps cleared* (true) while a subject reading the complex/control version would have judged the statement *A clever commander had to clear some important field camps* (false). The ratio of true to false statements was 1:1 across the entire experiment. Out of 56 possible cases (28 items times two conditions), 21 comprehension tests targeted the interpretation of the ellipsis. Other statements targeted either the antecedent or other parts of the stimulus sentences. Participants were given the opportunity to take a break after completing half of the experiment.

Predictions

If ellipsis is interpreted via a memory pointer mechanism (Frazier & Clifton, 2005; Martin & McElree, 2008) or, equivalently, a cost-free whole-clause copying mechanism (Frazier & Clifton, 2001), we expect no effect of the antecedent complexity at the critical region – that is, the ellipsis site – in the elided conditions. However, under the copying account of Murphy (1985), we expect longer reading times at the critical region for sentences with complex antecedents in the ellipsis conditions only. As no clause intervenes between antecedent and ellipsis site, Murphy’s theory predicts that readers should not fall back on a discourse-based processing mechanism, which would otherwise lead us to expect no effect of antecedent complexity. Finally, if more elaborate antecedents are easier to retrieve from memory, as would be expected given the findings of Hofmeister (2011), reading times at the critical region should be shorter for sentences with complex antecedents in the ellipsis conditions.

Note that both the Murphy (1985) and Hofmeister (2011) accounts predict an interaction between antecedent complexity and elision. This is important because antecedent complexity is completely confounded with the ellipsis site's position in the sentence. Any main effect of the complexity manipulation could thus be due to changes in participants' reading speed as they progress through the sentence (Demberg & Keller, 2008; Ferreira & Henderson, 1993). Martin and McElree (2008) circumvented this problem by adding material between antecedent and ellipsis in the simple antecedent conditions, which, however, increases the distance between the end of the antecedent clause and the ellipsis site, as well as introducing the possibility that the processing of the additional information may interfere with the encoding or retrieval of the antecedent. Being faced with two less-than-optimal alternatives, we opted to stay as close as possible to the designs of Frazier and Clifton (2000) and Murphy (1985), which did not keep sentence length constant across conditions.

Looking more closely at the results of Martin and McElree (2008), it should be noted that according to Foraker and McElree (2011), the failure to find an effect of a manipulation on processing speed in an SAT paradigm by itself does not entail that there should also be no effect on reading times in comparable self-paced reading or eye-tracking studies. Foraker and McElree (2011) argue that even if only the asymptotic accuracy – the highest level of accuracy that participants are able to reach with their grammaticality judgments – is affected in SAT, reading times in self-paced reading or eye-tracking may differ between conditions due to retrieval failures or low-quality interpretations. More specifically, a drop in asymptotic accuracy in SAT may translate to higher reading times due to reprocessing (McElree & Nordlie, 1999). Martin and McElree (2008) largely failed to find effects of antecedent complexity on asymptotic accuracy, with the exception of their Experiment 6, where an additional full noun phrase within the antecedent lowered accuracy. Based on this isolated result, higher reading times should be predicted for antecedents containing more full noun phrases. However, in Martin and McElree's other experiments, which also included an eye-tracking study, the presence of additional noun phrases in complex antecedents did not measurably affect accuracy or reading times, calling the result of Experiment 6 into question. We thus take Martin and McElree's evidence to point more strongly in the direction of there being no effect of antecedent complexity on ellipsis processing across paradigms, and indeed this appears to be the position adopted by the authors.

Data analysis

All data from the first participant were discarded before analysis as this session was considered a trial run, which revealed several minor mistakes. The remaining data were analyzed using the statistics software **R** (R Core Team, 2016). Linear mixed-effect models were fit to reading times and question response accuracies with the package **rstanarm** (Gabry & Goodrich, 2016), which provides an interface between **R** and the **Stan** programming language for Bayesian statistical inference (Stan Development Team, 2016). The data and code for both experiments will be released with the publication of this article.

One advantage of Bayesian inference in **Stan** is that a hierarchical linear model can almost always be fit with full variance-covariance matrices for subject and item random effects (Sorensen, Hohenstein, & Vasishth, 2016); this is often difficult to achieve with the **lme4** function (Bates, Mächler, Bolker, & Walker, 2015; see Bates, Kliegl, Vasishth, & Baayen, 2015 for further discussion). Another advantage is the more straightforward interpretation of results in a Bayesian setting. Instead of computing confidence intervals, which somewhat unintuitively refer to hypothetical repeated sampling (Hoekstra, Morey, Rouder, & Wagenmakers, 2014), a Bayesian credible interval specifies plausible values of the parameters given the data at hand. This makes inference much more straightforward compared to Null Hypothesis Significance Testing (see Nicenboim & Vasishth, 2016 for a review).

Reading times below 150 ms, which are (arguably) unlikely to be generated by linguistic processes, were removed prior to analysis; this resulted in a loss of less than 1% of data. The experimental factors were sum-coded. For the factor antecedent complexity, the complex conditions were coded as 1 and the simple conditions were coded as -1 , respectively. For the factor elision, the ellipsis conditions were coded as 1 and the control conditions were coded as -1 . As visual inspection of the reading time distributions suggested some amount of heteroscedasticity in the data, the Box-Cox procedure (Box & Cox, 1964) was applied, which suggested reciprocal transformation of reading times ($1/RT$) and logarithmic transformation of question-response times. Reciprocal reading times were multiplied by a factor of -1000 to allow for easier interpretation. All models included random intercepts and slopes by subjects as well as by items for each estimated parameter, including interaction parameters. The prior distribution for each estimated parameter was a normal distribution with mean zero and a standard deviation of 2.5, except for the intercept, for which a standard deviation of 10 was used. The LKJ prior (Lewandowski, Kurowicka, & Joe, 2009) with parameter 1 was

used for the variance-covariance matrices of the random effects for subjects and items;⁵ this imposes a regularization on the prior distribution of the variance-covariance matrix (see Stan Development Team, 2016 for details, and Sorensen, Hohenstein, & Vasishth, 2016 for a tutorial intended for psycholinguists). Besides fitting models to individual regions of interest, as is commonly done in psycholinguistics, we also fitted a model that took into consideration all data points from the second-to-last region leading up to the ellipsis site (crit−2) to the second region after the ellipsis site (crit+2). The region predictor was coded using a successive differences contrast, meaning that the model is estimating the *differences* in processing times between each two adjacent regions, starting with region crit−2. The region-by-region analyses can thus be seen as nested comparisons within the overall model (see Nicenboim, Logačev, Gattei, & Vasishth, 2016). To account for the fact that reading times within the same trial are not independent, we added a random intercept by trial to the model.

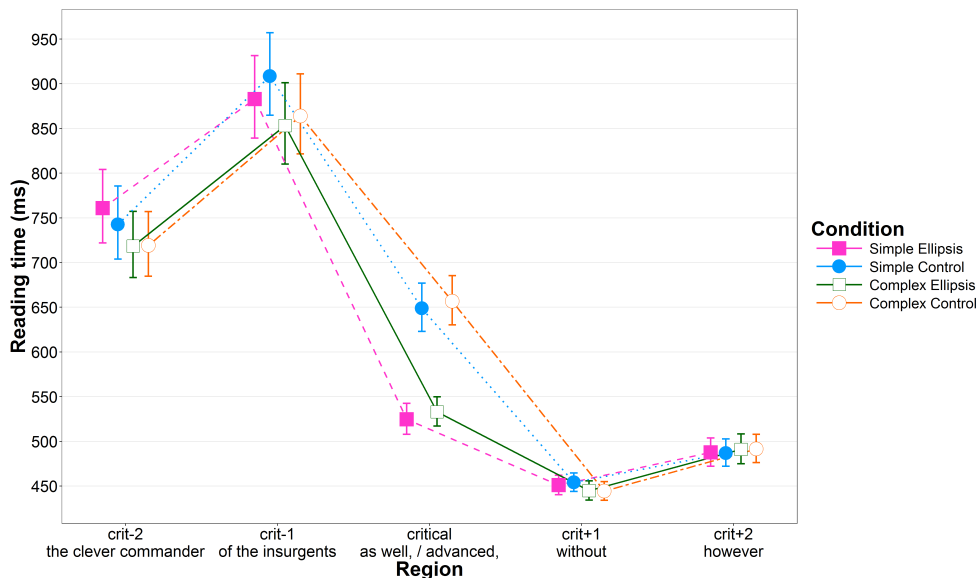
Four sampling chains with 4000 iterations each were run for each model, with a warm-up period of 2000 iterations. We report the estimated parameters, along with their 95% credible intervals and the posterior probability that the parameter’s true value is greater than zero. We judge there to be evidence for an effect if zero is not included in the associated 95% interval. We consider there to be weak evidence for an effect – which is to be distinguished from the effect itself being weak – if zero is within the interval but the probability of the parameter being above or below zero is high (> 95%).

Results

Question responses

Question response accuracy was 88% for all items and 85% for target items. The analysis of response accuracies revealed no effects of the experimental manipulations. However, there is some evidence of response times being elevated for ellipsis versus control sentences ($\hat{\beta}=0.031$, CrI: $[-0.004,0.067]$, $\Pr(\hat{\beta}>0)=0.96$). See Figure 3.2 for the results.

Figure 3.1: **Experiment 1 – Mean reading times (back-transformed, on reciprocal scale) by region and condition.** Error bars represent 95% confidence intervals.

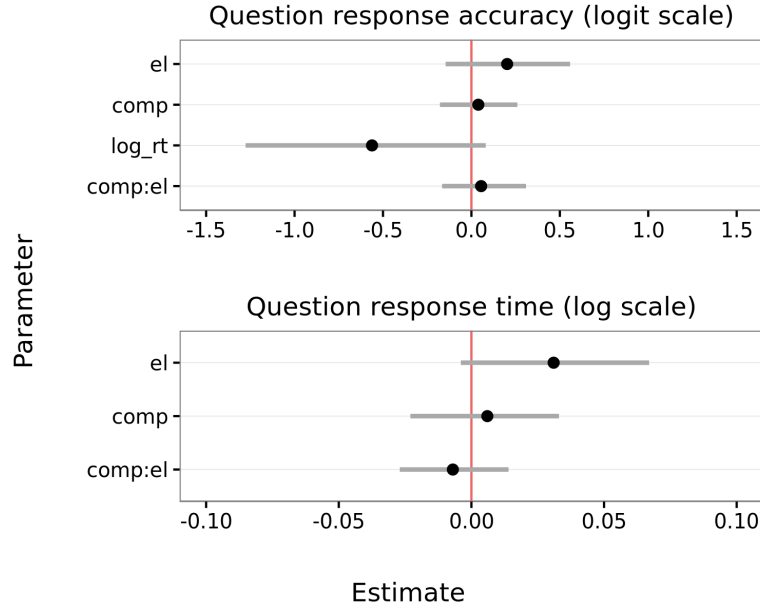


Reading times

Figure 3.1 plots back-transformed mean reciprocal reading times (see above) by region, along with 95 % confidence intervals. Figure 3.3 shows the results for individual regions of interest. Analyzing the regions between antecedent and ellipsis site is critically important because of the concern that processing spillover from the antecedent may influence reading times at the critical regions. In region crit–2, sentences with complex antecedents were read faster than sentences with simple antecedents ($\hat{\beta} = -0.029$, CrI: $[-0.056, -0.002]$, $\Pr(\hat{\beta} > 0) = 0.02$), indicating that there is no spillover from processing the antecedent. The complexity advantage – which may reflect a speedup due to having processed additional words – persists into region crit–1 ($\hat{\beta} = -0.025$, CrI: $[-0.049, 0.001]$, $\Pr(\hat{\beta} > 0) = 0.03$), but disappears as soon as the critical region is encountered. If anything, antecedent complexity *increased* reading times at the critical region, both for ellipsis and control sentences, but the evidence is very weak ($\hat{\beta} = 0.012$, CrI: $[-0.019, 0.043]$, $\Pr(\hat{\beta} > 0) = 0.78$). The critical region shows a main effect of elision, such that ellipsis is processed faster than the lexical

⁵An LKJ prior with parameter 1 assigns equal probability mass to all possible correlation values. An $sd(0, 10)$ prior on the intercept gives values between ± 19.6 on the transformed $-1000/RT$ scale, and values close to ± 1600 ms on the original scale. This uninformative prior serves as a sanity check to see if the model can recover a sensible intercept.

Figure 3.2: **Experiment 1 – Question response times and accuracies: caterpillar plot of means and 95% credible intervals for parameters of interest.** el = elision, comp = complexity, log_rt = log response time.

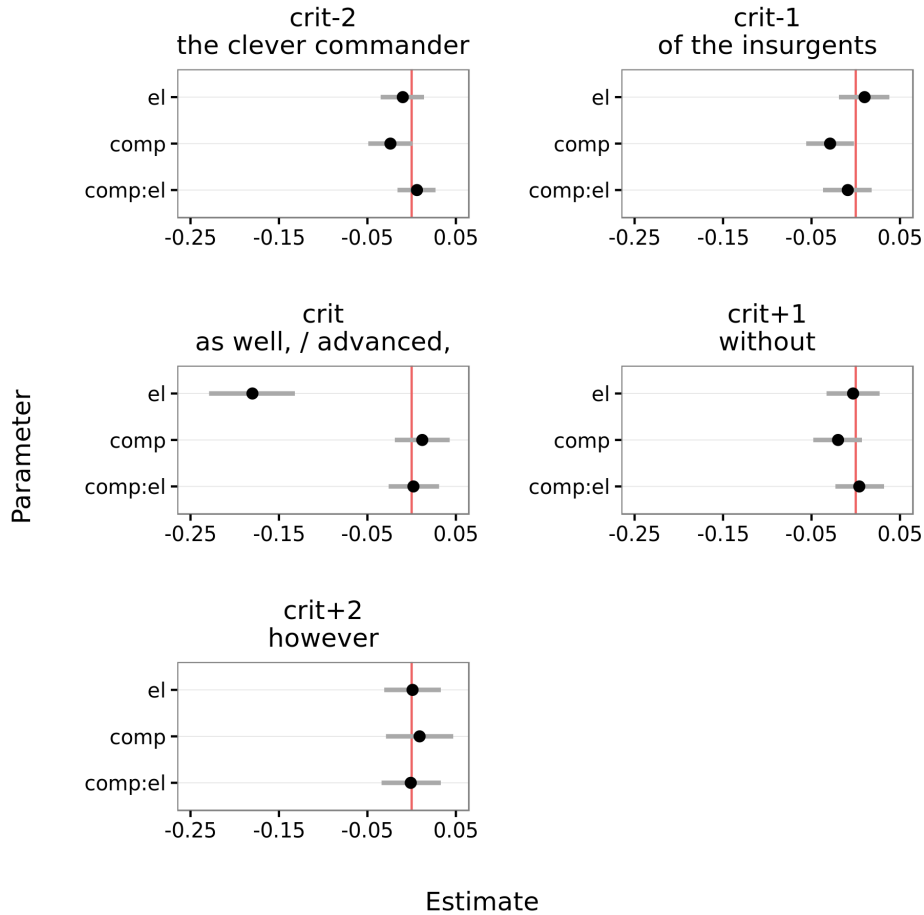


verbs in the control conditions ($\hat{\beta} = -0.180$, CrI: $[-0.229, -0.132]$, $\Pr(\hat{\beta} > 0) \approx 0$). In the region following the ellipsis (crit+1), there is again weak evidence for a complexity-induced speedup ($\hat{\beta} = -0.020$, CrI: $[-0.048, 0.007]$, $\Pr(\hat{\beta} > 0) = 0.08$). There was no evidence for an interaction between antecedent complexity and elision in any region.

Figure 3.4 shows the estimated parameters from the model fitted to regions crit–2 through crit+2 together. In addition to an overall complexity advantage ($\hat{\beta} = -0.029$, CrI: $[-0.054, -0.005]$, $\Pr(\hat{\beta} > 0) = 0.01$), sentences with complex antecedents show a smaller difference between region crit–1 and the critical region ($\hat{\beta} = 0.037$, CrI: $[0.003, 0.070]$, $\Pr(\hat{\beta} > 0) = 0.98$), but a larger difference between the critical region and region crit+1 ($\hat{\beta} = -0.032$, CrI: $[-0.065, 0.002]$, $\Pr(\hat{\beta} > 0) = 0.03$).⁶ This finding matches the patterns observed in the by-region analyses. Elision, meanwhile, increases the difference in processing times between regions crit–1 and the critical region ($\hat{\beta} = -0.170$, CrI: $[-0.212, -0.127]$, $\Pr(\hat{\beta} > 0) \approx 0$), but decreases the difference between the critical region and region crit+1 ($\hat{\beta} = 0.176$, CrI: $[0.133, 0.221]$, $\Pr(\hat{\beta} > 0) \approx 1$), matching the main effect of elision observed in the critical region itself.

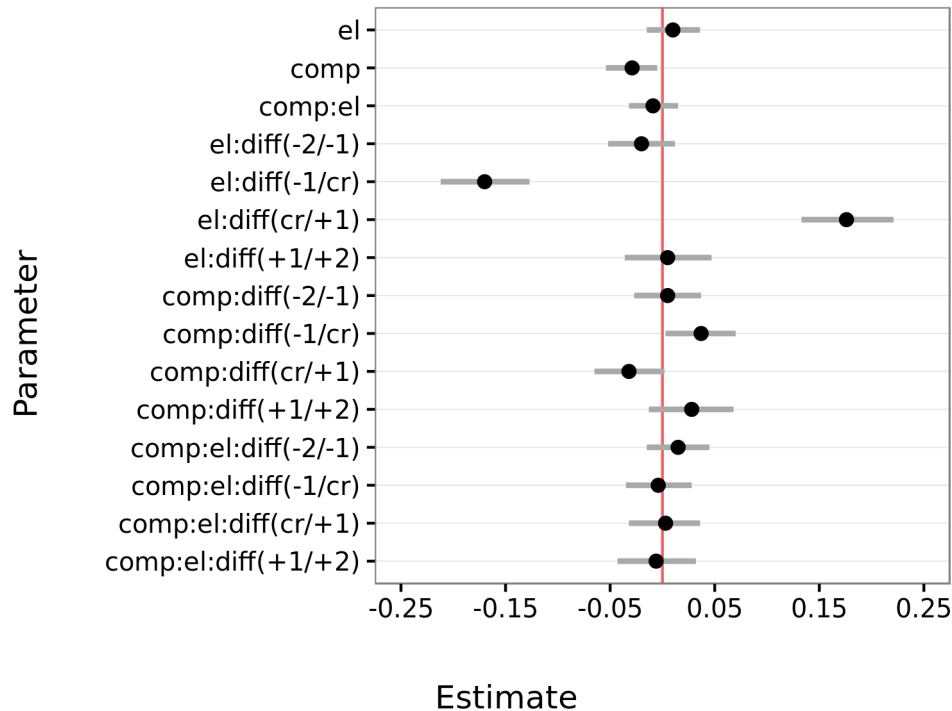
⁶Both difference parameters are negative, which means that an interaction with a negative sign indicates a larger difference while one with a positive sign indicates a smaller difference.

Figure 3.3: **Experiment 1 – Caterpillar plot of means and 95% credible intervals for parameters of interest, separate models fit across regions crit–2 through crit+2.** el = elision, comp = complexity.



While there was no evidence of an interaction between the elision and complexity manipulations at the critical region, the main effect of complexity that is visible in the overall analysis is of theoretical interest. As we were interested in further investigating the effect of antecedent complexity on reading times for the critical region, we subjected the relevant coefficient from the single-region model, whose estimate showed only very weak evidence for being positive, to a Bayes Factor analysis. A hypothesis test based on the Bayes Factor provides a way to quantify the support for the model under which the observed data are most likely (Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010). We chose to perform multiple order-restricted analyses, meaning that sampling was restricted to either only positive or only negative coefficient values, respectively, in order to better gauge the amount

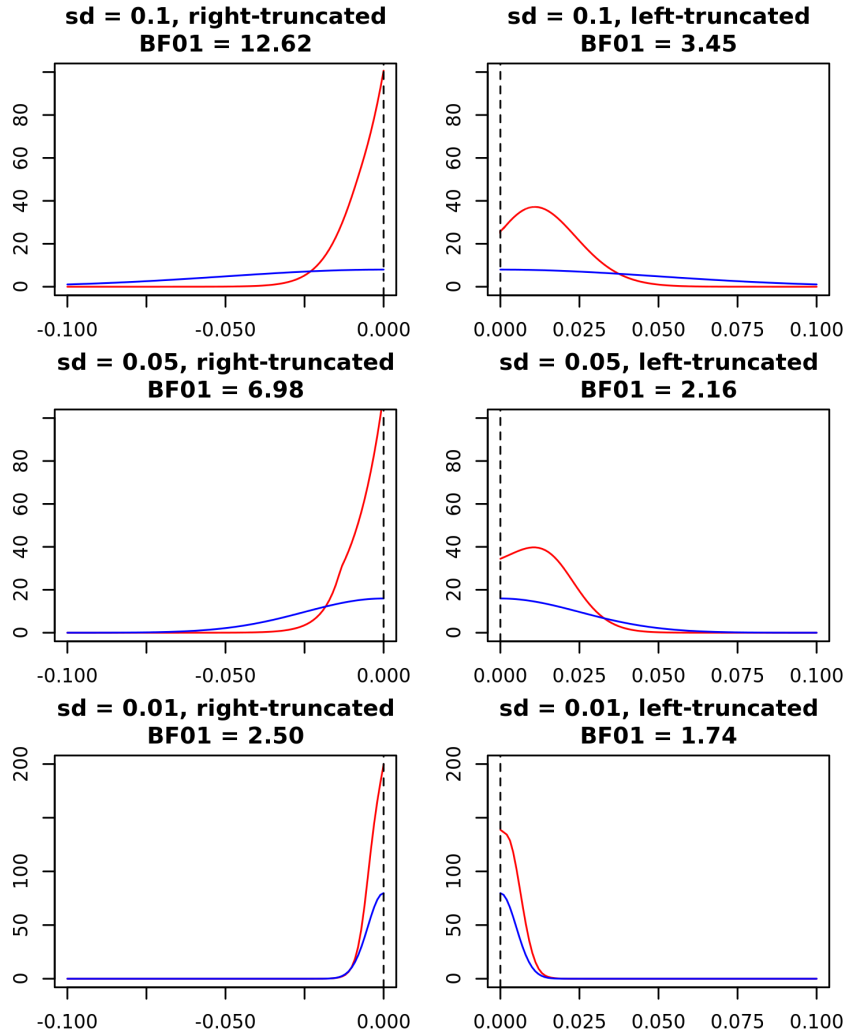
Figure 3.4: **Experiment 1 – Caterpillar plot of means and 95% credible intervals for parameters of interest, one model fit across regions crit–2 through crit+2.** *diff* parameters represent successive differences between regions, in linear order. el = elision, comp = complexity.



of evidence for or against the coefficient in question being different from zero in a given direction. By using left- or right-truncated prior distributions, it is possible to quantify support against the null hypothesis and also in favor of it, relative to a directed alternative hypothesis. Additionally, the conclusions from the Bayes Factor test are more conservative than those based on credible intervals: Even a 95% credible interval that does not include zero does not guarantee a high value of the Bayes Factor, that is, it does not guarantee strong support for the alternative hypothesis (Wagenmakers et al., 2010).

We used the Savage–Dickey density ratio method (J. M. Dickey & Lientz, 1970) to compute the Bayes Factor, following Wagenmakers et al. (2010). Even though the posterior distributions for the model parameters are generally not sensitive to the prior settings, the Bayes Factor is acutely so. When priors are too wide (too uninformative), the alternative hypothesis assigns too much prior mass to values that yield very low likelihoods. This in turn means that without proper specification of priors, the null hypothesis would be always more

Figure 3.5: **Experiment 1 – Density plots for the Bayes factor analysis.** Blue = prior density, red = posterior sample density.



likely than the alternative hypothesis, since its prior mass is concentrated in zero. Three normal distributions of different widths were used as priors on the complexity coefficient (on the transformed scale): $N(0, 0.05)$, $N(0, 0.025)$ and $N(0, 0.005)$. Model specifications were otherwise left unchanged. With $N(0, 0.05)$, 95% of the probability mass will be in the interval $[0, 0.1]$ for a left-truncated normal distribution and in the interval $[-0.1, 0]$ for a right-truncated normal distribution. The corresponding values on the original scale are 69 ms, 34

ms and 7 ms, respectively. Figure 3.5 shows the calculated Bayes factor values for the different priors, along with density plots for the prior distributions versus posterior samples.⁷

For the right-truncated priors, the value of the Bayes factor depends heavily on the spread of the distribution: For the widest prior, the null hypothesis is more than twelve times as likely as the alternative hypothesis that the complexity effect is negative, while for the narrowest prior it is still between two and three times as likely. For the left-truncated priors, the null hypothesis is between two and three-and-a-half times as likely to be true, depending on the spread – and thus the informativeness – of the prior. Note that unlike for left-truncated priors, the point of maximum probability density for the posterior samples given right-truncated priors is always at zero. On the whole, the analysis shows that for all but the most narrow distributions the prior restriction that the complexity effect should be negative or null leads to more evidence in favor of the null hypothesis compared to when the effect is restricted to be positive or null. There is thus evidence that the effect is probably not negative, and more likely to be null than positive, even though the latter conclusion is only weakly supported if one adheres to common interpretation guidelines for the Bayes factor (Raftery, 1995).

Discussion

Three main results were obtained in the current study:

- (I) Ellipsis was processed faster than the lexical verbs used in the control conditions.
- (II) Overall, having processed a longer and more complex antecedent led to faster reading times across later regions.
- (III) At the critical region, the speedup was temporarily suspended. An analysis based on the Bayes factor yielded evidence in favor of a null effect of the complexity manipulation at the critical region.

Finding (I) may be trivially explained by the fact that the critical region was shorter in the ellipsis compared to the control conditions for most items. The prediction of an interaction between the elision manipulation and antecedent complexity was not borne out in the data, a result that is most consistent with pointer-based accounts of ellipsis resolution (Frazier & Clifton, 2001, 2005; Martin & McElree, 2008). Furthermore, finding (III) suggests that the overall speedup induced by the complexity manipulation was nullified at the critical region,

⁷Note that as we are using the logspline package (Koopberg, 2016) in R to estimate the marginal posterior density at zero, there is some variation in the calculated Bayes factors when the computations are run multiple times. In our test runs, each Bayes factor varied within the limits of approximately ± 0.5 of the values shown in Figure 3.5.

rather than turning into a slowdown, further supporting pointer-based approaches. Even if there had been evidence of such a slowdown – which, given the Bayes factor results, would need to be of a very small magnitude – we would still have needed to explain why it would appear in both the ellipsis and control conditions (see discussion below).

We believe that finding (II) has a mechanistic explanation: readers tend to read faster if they are deeper into the sentence already. The ‘complexity’-induced speedup would thus be a length or, equivalently, a linear position effect. Given our initial predictions, we found no evidence that increased antecedent complexity slowed down the processing of the ellipsis, as predicted by Murphy’s (1985) account, or contrariwise led to speedier processing of the ellipsis, as predicted by the account of Hofmeister (2011). Rather, the Bayes factor analysis showed that the data are largely inconsistent with the assumption that increased complexity leads to faster processing.

A potential influence of within-sentence wrap-up

The results at the critical region warrant closer inspection, as one might argue that the observed temporary suspension of the speedup effect could be due to wrap-up caused by the comma at the end of the second conjunct. If wrap-up reflects integration processes at the end of a clause, since integrating more complex meanings takes longer, readers will possibly spend more time on the final region of the second conjunct if the first conjunct contains more information. This may then momentarily cancel out the speedup that is visible before and after the critical region. When designing the experiment, we avoided having the ellipsis followed by a period, neglecting that commas also create wrap-up effects, albeit of a smaller magnitude (Warren, White, & Reichle, 2009). In our defense, it is quite impossible to study clausal ellipsis without having the end of the elided clause marked somehow in the input. In any case, there is evidence from eye-tracking suggesting that wrap-up at punctuation marks such as commas is not influenced by the complexity of the sentence (Rayner, Kambe, & Duffy, 2000; Warren et al., 2009), which casts doubt on the assumption the complexity effect observed at the critical region is only due to the comma.

If anything, one would need to claim that the position-based speedup in reading that has been observed repeatedly (Demberg & Keller, 2008; Ferreira & Henderson, 1993) is completely suspended during wrap-up. As a quick check of this assumption, we fitted a Bayesian linear mixed-effects model to the data from our filler items. In this model, linear position of the presentation region within the sentence and the presence or absence of a comma were used as predictors. The comma factor was sum-coded with comma present

being coded as 1, and region position was entered as a continuous predictor. The model revealed that there was indeed a position-related speedup ($\hat{\beta} = -0.010$, CrI: $[-0.021, 0.000]$, $\Pr(\hat{\beta} > 0) = 0.027$), as well as a comma-induced slowdown ($\hat{\beta} = 0.057$, CrI: $[0.025, 0.091]$, $\Pr(\hat{\beta} > 0) \approx 1$), and an interaction with a negative sign: the speedup effect appears to be *stronger* rather than weaker when a comma is present ($\hat{\beta} = -0.008$, CrI: $[-0.017, 0.001]$, $\Pr(\hat{\beta} > 0) = 0.038$). This implies that the presence of a comma probably did not result in a suspension of the speedup effect observed in our experimental items.

Given these findings, the possibility arises that the speedup was still in effect at the critical region, but was counteracted by a complexity-induced slowdown in the vein of Murphy (1985), resulting in the two effects canceling each other out. Under this assumption, however, one is left asking why the slowdown should also be present in the control conditions.

A possible issue of parallelism

There may be other reasons for not expecting an effect of the manipulation in our materials. Particularly, our use of the conjunction *and*, ‘and’, might be critical to understanding our failure to observe an interaction between antecedent complexity and ellipsis processing. The results of a cross-modal priming study by Callahan, Shapiro, and Love (2010) are informative in this regard. In their Experiment 2, Callahan et al. presented sentences like (17) auditorily. Words that were either related or unrelated to the verb *read* in the initial clause (related: *reviewed*, unrelated: *reserved*) appeared on the screen at the positions marked in the example. Subjects were required to read these words aloud.

- (17) The doctor read the chart of the child with the broken arm [1] during his morning rounds, and [2] the insurance agent in [3] the tacky suit did as well [4] in order to become more familiar with the case.

Results showed that naming responses to related words were faster at probe positions 3 and 4, but not at positions 1 and 2. Experiment 1 used only probe positions 1 and 2, revealing a priming effect at position 2, but not at position 1. Despite the priming effect at the conjunction itself not appearing consistently, Callahan et al. (2010) conclude that material from the first clause is reactivated during the processing of the second clause. The conjunction *and* arguably induces an expectation of parallelism, causing the retrieval and subsequent maintenance of the verb *read*, or possibly of the entire associated proposition, allowing for easier integration with the second conjunct. Callahan et al. (2010) suggest that active maintenance of antecedent information may be achieved through repeated retrievals prior to the ellipsis site which are

cued by the conjunction. Even though parallelism has long since been known to facilitate the processing of coordinate structures (Frazier, Munn, & Clifton Jr., 2000; Frazier, Taft, Roeper, Clifton Jr., & Ehrlich, 1984), Callahan et al. (2010)'s *sustained reactivation hypothesis* is, to our knowledge, the first account to explicitly link this observation to working memory.

If the presence *und*, 'and', in our stimuli led participants to assume parallelism between the conjuncts, causing them to actively maintain information from the antecedent clause, there is an alternative explanation for the prolonged speedup effect we observed: participants were simply eager to reach the end of the second conjunct, since this is the point where the two propositions can be integrated. Crucially, sustained reactivation also obviates the need for a laborious retrieval at the critical region, since the necessary information is already available, thus predicting no detrimental effect of the complexity manipulation, apart from possible costs associated with discourse integration.

Even without sustained reactivation being a factor our stimuli, the lack of an interaction between antecedent complexity and elision can be explained if one assumes that lexical verbs can also trigger retrievals. This might be true especially in coordinate structures, where parallelism reinforces the semantic association between the conjuncts. Indeed, the control conditions in many of our sentences imply a causal connection between the two propositions, such as the commander advancing after the enemy's field camps have been cleared in (18).

- (18) The army cleared some important field camps and the clever commander of the insurgents advanced.

Pointer-based approaches (Frazier & Clifton, 2001; Martin & McElree, 2008) can account for the result by claiming that retrieval time is negligible across conditions, and that any complexity-induced slowdown reflects integration difficulty *after* retrieval. However, one would then need to assume that this integration difficulty is limited to *and*-conjoined sentences like the ones used in our study: Frazier and Clifton (2001) found no complexity effect for two-sentence discourses – but recall the study's limitations noted in the introduction – and Martin and McElree (2008) found no complexity-induced change in ellipsis processing times for *but*-conjoined sentences.

A more precise notion of complexity-based facilitation is needed

Assuming that retrieval takes place in both the ellipsis and control conditions, the observed processing pattern would be more in line with the reasoning of Murphy (1985), where it takes more time to copy more information from the antecedent, than with that of Hofmeister (2011), where elaboration should lead to facilitation. Indeed, our analyses showed more evidence for the former view than the latter. However, as was pointed out before, (Hofmeister, 2011, p. 395) assumes that not all kinds of elaboration aid retrieval; only strongly associated features of a memory trace are predicted to have a facilitatory effect. Uncommon feature combinations (*lovable dictator*), while increasing encoding time, will impede retrieval instead of providing easier access to the target. While Hofmeister's results show that there is no direct connection between encoding time and retrieval time, it is by no means clear whether the elaboration provided by the complexity manipulation in the current study should have yielded any additional facilitatory anchors for memory access. The answer would depend, among other things, on whether the component parts of the antecedent are visible to the retrieval probe. If we assume that the search process that is initiated when a clausal ellipsis is encountered focuses on finding a phrase containing a verb, which is the semantic core of a clause, it might ignore any adjuncts or auxiliaries attached to it. If the search process is serial, the presence of such elements may even result in longer processing times. Taken at face value, however, the theory of Hofmeister (2011) should predict facilitation for our stimuli, given that clausal adjuncts are to sentence meaning what adjectives, as used by Hofmeister, are to a noun phrase, that is, elaborative modifiers. Thus, if the presence of an adjective influences the retrieval process, so should the presence of a clausal adjunct.

Conclusion

In short, Experiment 1 showed evidence in favor of a null effect of antecedent complexity on ellipsis processing times. The results should, however, be interpreted with a certain amount of caution. On the methodological side, one important shortcoming is that the experiment used sentences conjoined by *and*, 'and', possibly causing the control conditions not to work as intended. Our second study sidesteps the issue of parallelism by using *but*- instead of *and*-coordinated sentences. Unlike *and*, *but* evokes no expectation of parallelism between the two conjuncts, and indeed parallelism does not facilitate processing for *but*-conjoined

sentences (Knoeferle, 2014). The main goal of Experiment 2 was to investigate whether antecedent complexity effects in ellipsis processing are sensitive to task demands, as suggested by Phillips and Parker (2014) and Paape (2016). The design is inspired mainly by Swets et al. (2008)'s investigation of parsing preferences for a temporary syntactic ambiguity.

Experiment 2

Drawing from the literature on 'good-enough' processing (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003), Swets et al. (2008) explored whether asking different kinds of comprehension questions would influence readers' on-line processing of syntactically ambiguous sentences in a self-paced reading experiment. The construction in question involves a relative clause whose attachment site is initially not obvious, as shown in (19). The gender of the reflexive *himself/herself* disambiguates the structure towards attachment to either the first NP (N1) or the second (N2) in (19b,c), but not in (19a).

- | | | | |
|------|----|--|----------------------|
| (19) | a. | The maid of the princess who scratched herself . . . | ambiguous |
| | b. | The son of the princess who scratched himself . . . | N1 attachment |
| | c. | The son of the princess who scratched herself . . . | N2 attachment |
| | | . . . in public was terribly humiliated. | |

Subjects were divided into three groups according to the kind and frequency of comprehension questions that appeared along with the experimental sentences. One group of participants was asked questions that targeted the interpretation of the relative clause, such as *Did the maid scratch in public?* A second group answered questions that did not target the relative clause, and indeed did not require much attention to the sentences' contents, such as *Was anyone humiliated?* A third group was also asked these superficial questions, but only on one out of every twelve trials.

Swets et al. (2008) found that participants expecting questions about the relative clause attachment took longer to read the post-disambiguation region if the attachment had been disambiguated toward N1 (19b) than for both of the other conditions. The pattern for readers in the two other groups looked different: they were faster in the ambiguous condition (19a) than in both the N1 and N2 conditions. These results indicate that readers' syntactic processing strategies may change according to task demands. If participants know that their interpretation of an ambiguous sentence will be probed, they appear to preferentially choose one possibility, namely N2 attachment. If, however, participants do not have to worry about

their interpretation being queried explicitly, they enjoy a processing advantage due to the possibility of not making an attachment decision at all. This is commonly referred to as underspecification.

Given that effects of task demands have also been observed in discourse processing (Foertsch & Gernsbacher, 1994), it is conceivable that people have more than one strategy available for the resolution of ellipsis. Another possibility is that readers can be somewhat selective in terms of what information they retrieve – or, alternatively, maintain and integrate – at the ellipsis site. If comprehension of the elliptical clause is not probed too deeply, they might even opt to not resolve the anaphor at all. This latter view is rather extreme, given the implication that readers never make an effort to understand experimental stimuli unless explicitly motivated to do so. One might argue that since reliable effects of experimental manipulations can be observed even in studies which feature no or only shallow tests of comprehension, there must be some intrinsic motivation to interpret sentences even when there is no payoff. While this is a valid point, it is by no means clear whether we can rely on the compliance of our subjects in all cases, especially in light of recent findings on ‘good-enough’ processing.

While Experiment 1 investigated bare argument ellipsis (‘stripping’) in German, Experiment 2 used English VP ellipsis constructions, much like the aforementioned studies of Frazier and Clifton (2000) and Martin and McElree (2008). As the discussion of Experiment 1 suggested, the control conditions used in the previous study may not have served their purpose as intended, so for Experiment 2 we dispensed with them. Instead, subjects were divided into two groups which received different kinds of comprehension probes during the experiment. Since many of these were directly related to the interpretation of the elided VP, we can assume that any group-specific effects we observe will be connected to the presence of ellipsis, rather than to other aspects of the stimuli.

Materials

A sample stimulus from Experiment 2, along with two of the associated comprehension probes, is shown in (20). As before, diamonds indicate the boundaries of presentation regions during the experiment. The experimental factors used in this study were antecedent complexity (simple vs. complex) and probe type (superficial vs. detailed). In the current study, simple antecedent clauses always contained only a simple object NP (see below), while in complex antecedent clauses this object NP in turn contained a genitive modifier as well as additional adjectives. Note that unlike in Experiment 1, the antecedent complexity manipulation did not

change the number of presentation regions. Probe type remained constant throughout each experimental session and divided subjects into two groups. The study thus employed a 2 (within-subjects) \times 2 (between-subjects) design. A total of 36 items and 160 fillers were presented in random order during each experimental session. The stimuli are listed in Appendix III.

(20) **Antecedent preamble**

The advanced students \diamond loved \diamond . . .

Simple antecedent

. . . the afternoon session, \diamond . . .

Complex antecedent

. . . the late afternoon session's many illustrative examples, \diamond . . .

Continuation

. . . but \diamond as of late \diamond it was evident \diamond that \diamond the mathematics lecturer \diamond did not,
 \diamond as \diamond the time-consuming preparation \diamond really \diamond exhausted her.

Superficial probe

A mathematics lecturer was mentioned.

Detailed probe

A lecturer did not love an afternoon session's examples.

An additional difference in comparison to Experiment 1 is the presence of an extra clausal layer between antecedent and ellipsis. This increases the distance between the loci of encoding and integration of the antecedent, and may make subjects less likely to adopt a strategy based on memory maintenance or 'sustained reactivation' as observed by Callahan et al. (2010).

A negation occurred as part of the critical ellipsis region (*did not*) in half of the experimental items, like in (20). For the other half, the negation instead occurred in the antecedent region (*The advanced students did not love . . .*) and the critical region consisted only of the auxiliary *did*. Comprehension probes appeared after each sentence in both groups, with equal numbers of true and false statements. As in Experiment 1, subjects were required to assess the veracity of the statements given the information in the preceding sentence.

Probes in the superficial group always followed the template ___ *was mentioned*, featuring either an entity that appeared in the preceding sentence or an unrelated entity that had not been mentioned. Probes in the detailed group randomly targeted either the antecedent or the ellipsis, with either unchanged or reversed polarity, and sometimes with parts of the original string replaced by novel terms (*Some students loved a morning session*; correct answer: false). Other aspects of the sentences were never targeted.

Participants

Eighty-one native speakers of English recruited from the University of Massachusetts, Amherst participated in the study. Forty-one subjects were assigned to the superficial probe group, the remaining forty to the detailed probe group. All subjects received course credit for their participation, and informed consent was obtained before each experimental session. The study was approved by the local Institutional Review Board of the Linguistics Department at the University of Massachusetts, Amherst.

Procedure

The procedure was largely the same as in Experiment 1, apart from the changes to the comprehension probes described above. Instead of masked self-paced reading, as described for Experiment 1, Experiment 2 used centered self-paced reading to avoid line breaks occurring inside the antecedent region. In centered self-paced reading, each region is presented in the center of the screen and replaced with the next region when the space bar is pressed. A fixation cross was presented for 1000 ms before each trial to mark the position of a given region's first character.

Predictions

Assuming that the overall speedup in the complex antecedent conditions of Experiment 1 was due to the use of *and*, which creates an expectation of parallelism, we should see no such effect in the current experiment, given that *but* was used instead. Should such an effect nevertheless appear, one would need to adopt a more task-oriented explanation, such as readers being anxious to get to the end of the sentence as quickly as possible. This kind of strategy might make sense if readers are afraid they might forget the information they need to answer the comprehension

questions. It would then also make sense for readers in the detailed probe group to show a larger speedup, as they can expect to be queried about the sentences' contents more rigorously.

If a parallelism requirement induced by *and* was responsible for masking any antecedent complexity effects related exclusively to ellipsis processing in Experiment 1, reading times at the ellipsis site in the current study may increase, decrease or be unaffected as the antecedent becomes more complex. The first possibility would be consistent with the predictions of Murphy (1985), unless the increased distance between antecedent and ellipsis in comparison to Experiment 1 (see materials section) causes subjects to fall back on 'discourse-based' processing. A decrease in ellipsis processing time for complex antecedents would support the notion of elaboration-based facilitation along the lines of Hofmeister's (2011) account. A null result, meanwhile, would lend credibility to approaches in which antecedent complexity is not expected to influence ellipsis processing at all (Frazier & Clifton, 2001, 2005; Martin & McElree, 2008).

On the other hand, the latter account would be called into question most strongly if the detailed probe group showed evidence of complexity effects at the point of retrieval while the superficial probe group did not. This would imply that task effects are a factor in ellipsis processing, and that the studies of Frazier and Clifton (2000) and Martin and McElree (2008) may have yielded null results due to subjects being insufficiently motivated to interpret sentences carefully.

Data analysis

Data analysis was carried out in a manner analogous to Experiment 1. The experimental factors antecedent complexity and probe type were sum-coded, with the levels 'simple' and 'superficial' receiving the value -1 and the levels 'complex' and 'detailed' receiving the value 1 , respectively. Again, all models featured the maximal random effects structure, to the exclusion of a random slope for probe type by subject, since this was a between-subjects factor. As before, models were fit to individual regions of interest as well as to all the data from within two regions around the ellipsis site together.

Results

Question responses

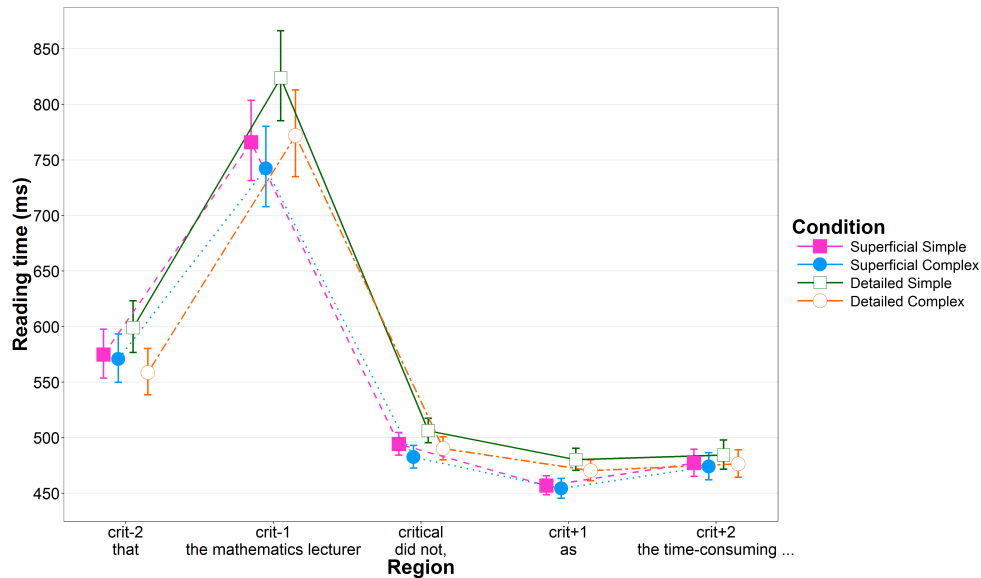
Mean question response accuracy across groups and conditions was 82%. Subjects in the superficial probe group reached 90% accuracy in both conditions, while subjects in the detailed probe group reached a mean accuracy of 74% in the simple antecedent condition, as compared to 72% in the complex antecedent condition. Results show that both increased antecedent complexity ($\hat{\beta} = 0.036$, CrI: [0.018, 0.053], $\Pr(\hat{\beta} > 0) \approx 1$) and membership in the detailed probe group ($\hat{\beta} = 0.23$, CrI: [0.176, 0.295], $\Pr(\hat{\beta} > 0) \approx 1$) increased response time. There was also an interaction between the factors ($\hat{\beta} = 0.024$, CrI: [0.006, 0.044], $\Pr(\hat{\beta} > 0) = 0.99$), such that antecedent complexity led to elevated response times only in the detailed group (nested comparisons: $\hat{\beta} = 0.06$, CrI: [0.036, 0.085], $\Pr(\hat{\beta} > 0) \approx 1$). Response accuracy was lower for the detailed group ($\hat{\beta} = -0.428$, CrI: [-0.667, -0.199], $\Pr(\hat{\beta} > 0) \approx 0$) and dropped with increased response time ($\hat{\beta} = -0.644$, CrI: [-1.000, -0.306], $\Pr(\hat{\beta} > 0) \approx 0$). The results are shown in Figure 3.7.

Reading times

Figure 3.6 plots back-transformed mean reciprocal reading times (see above) by region, along with 95% confidence intervals. Figure 3.8 shows the results for individual regions of interest. At region crit-2, sentences containing complex antecedents are read more quickly ($\hat{\beta} = -0.032$, CrI: [-0.056, -0.008], $\Pr(\hat{\beta} > 0) \approx 0$), an effect which persists throughout the subsequent regions (crit-1: $\hat{\beta} = -0.031$, CrI: [-0.057, -0.007], $\Pr(\hat{\beta} > 0) = 0.01$; critical region: $\hat{\beta} = -0.029$, CrI: [-0.049, -0.009], $\Pr(\hat{\beta} > 0) \approx 0$), until region crit+1, where there is only weak evidence of a speedup ($\hat{\beta} = -0.014$, CrI: [-0.036, 0.007], $\Pr(\hat{\beta} > 0) = 0.10$).

The combined analysis of reading times between regions crit-2 and crit+2 shows an overall facilitatory effect of complexity ($\hat{\beta} = -0.031$, CrI: [-0.052, -0.010], $\Pr(\hat{\beta} > 0) \approx 0$), along with an interaction with probe type ($\hat{\beta} = -0.027$, CrI: [-0.049, -0.005], $\Pr(\hat{\beta} > 0) = 0.01$) such that the speedup associated with more complex antecedents was larger in the detailed probe group. Probe type also influenced the changes in reading times between the critical region and region crit+1 ($\hat{\beta} = 0.026$, CrI: [-0.002, 0.054], $\Pr(\hat{\beta} > 0) = 0.96$), as well as between regions crit+1 and crit+2 ($\hat{\beta} = 0.037$, CrI: [-0.068, -0.006], $\Pr(\hat{\beta} > 0) = 0.01$): The negative difference between the critical region and region crit+1 was larger in the superficial group, as was the positive difference between regions crit+1 and crit+2. Model output is shown in Figure 3.9.

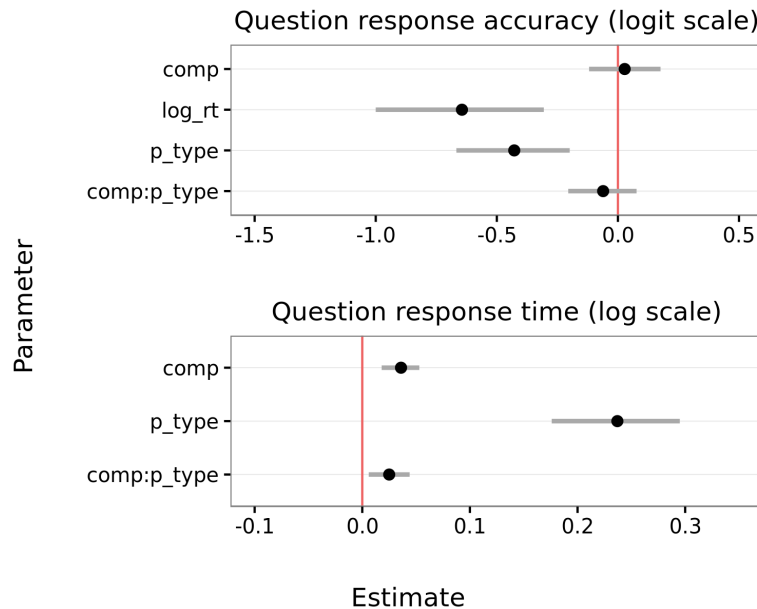
Figure 3.6: **Experiment 2 – Mean reading times (back-transformed, on reciprocal scale) by region and condition.** Error bars represent 95% confidence intervals.



As for Experiment 1, we conducted an additional analysis based on the Bayes factor, using the same procedure as before. For the current experiment, we were particularly interested in the interaction term of the model fitted at region crit+1; this is where the probe type manipulation had an effect on reading times, but the complexity manipulation did not appear to affect processing any differently than in the other regions. The lack of a differential influence is visible in the credible intervals of the three-way interactions with the region predictor in Figure 3.9, which are centered roughly around zero.

Figure 3.10 shows the results of the Bayes factor analysis. Predictably, left-truncated prior distributions yield evidence in favor of the null hypothesis, leading to the conclusion that the sign of the interaction term is very unlikely to be positive, contra Murphy (1985). While the null hypothesis is still favored with right-truncated prior distributions, the evidence is very weak: for the two tighter priors, it is not even two times as likely as the alternative. Therefore, it is possible that the greater overall speedup for sentences with complex antecedents that is visible in the detailed probe group affects region crit+1 just like the rest of the sentence. The Bayes factor results thus yield ancillary evidence that the probe type manipulation did not interact with the antecedent complexity manipulation in a way that would support either Murphy (1985) or Hofmeister (2011), given that the interaction is either null or otherwise not limited to the predicted region.

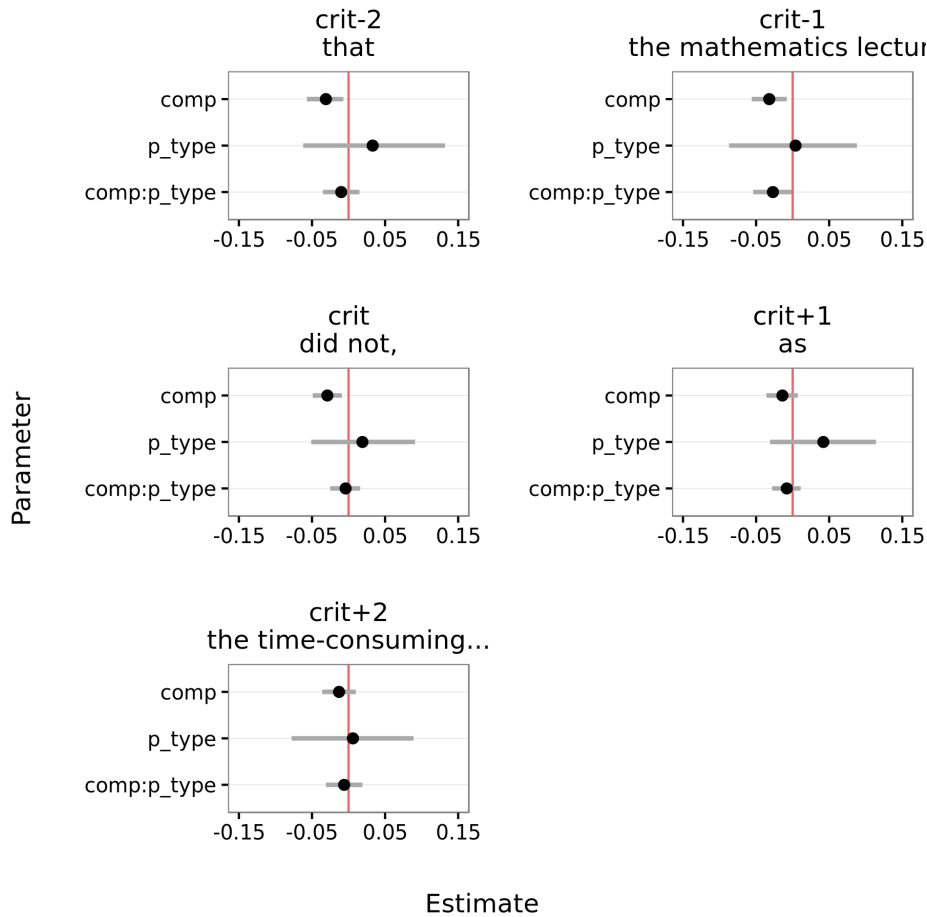
Figure 3.7: **Experiment 2 – Question response times and accuracies: caterpillar plot of means and 95% credible intervals for parameters of interest.** comp = complexity, p_type = probe type, log_rt = log response time.



Discussion

Given the results for the comprehension probe responses, we feel confident in claiming that our between-groups manipulation worked as intended: participants in the detailed probe group took longer to give an answer, and disproportionately longer than participants in the superficial group when the probe targeted a complex ellipsis antecedent. It thus appears that the detailed probes were indeed more difficult to answer, and that responding correctly became more difficult if information about either a more complex antecedent or a more complex ellipsis meaning was queried. However, we found no evidence of an interaction between probe type and antecedent complexity that would have been limited to critical ellipsis region. Assuming that participants in the detailed probe group processed the experimental stimuli more deeply, this result implies that the failure of earlier studies to find effects of antecedent complexity on ellipsis processing probably was not due to subjects' tendency to engage in 'good enough' processing. The findings of Experiment 2 are thus in line with the predictions of pointer-based approaches, and most strongly undercut those of Murphy (1985): under Murphy's account, subjects in the detailed group would have been expected to experience a greater slowdown due to increased antecedent complexity in the critical region, given the

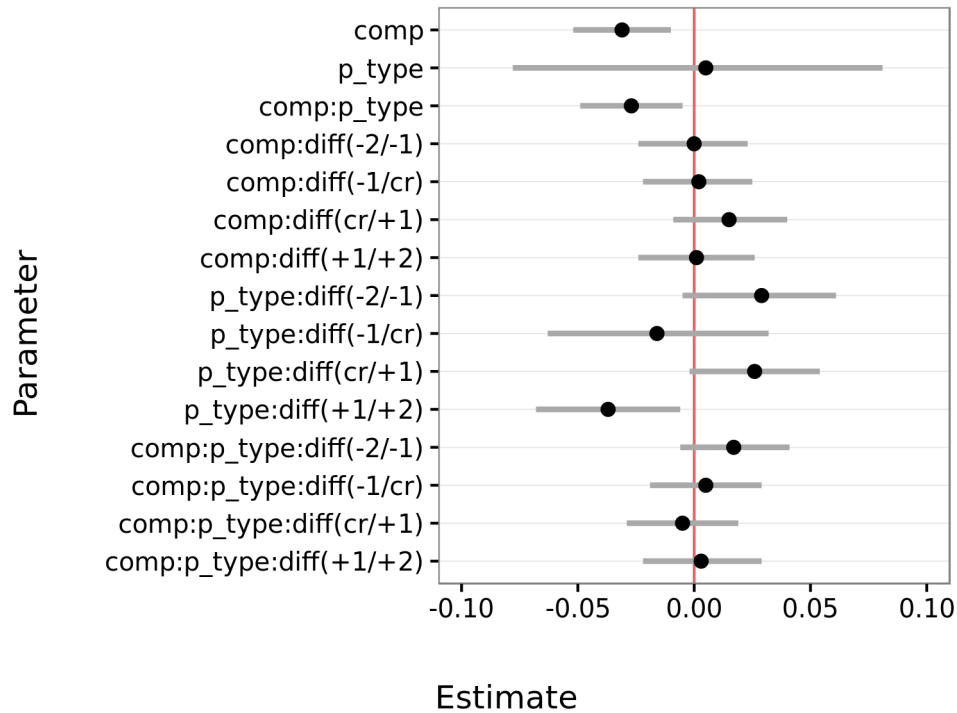
Figure 3.8: **Experiment 2 – Caterpillar plot of means and 95% credible intervals for parameters of interest, separate models fit across regions crit–2 through crit+2.** comp = complexity, p_type = probe type.



assumption that earlier null results were due to superficial processing. We also found no evidence that would have supported the account of Hofmeister (2011), given that there was no indication of speedier retrieval of complex antecedents within as well as across groups.

As in Experiment 1, having read a more complex antecedent was associated with faster reading times for later regions. For all regions of interest taken together, the speedup interacted with the probe type manipulation, such that the reduction in overall reading times was greater for the detailed probe group. This might indicate that members of the detailed probe type were more busy trying to remember the contents of complex antecedents, and thus withdrew resources from processing. We return to this point in the general discussion.

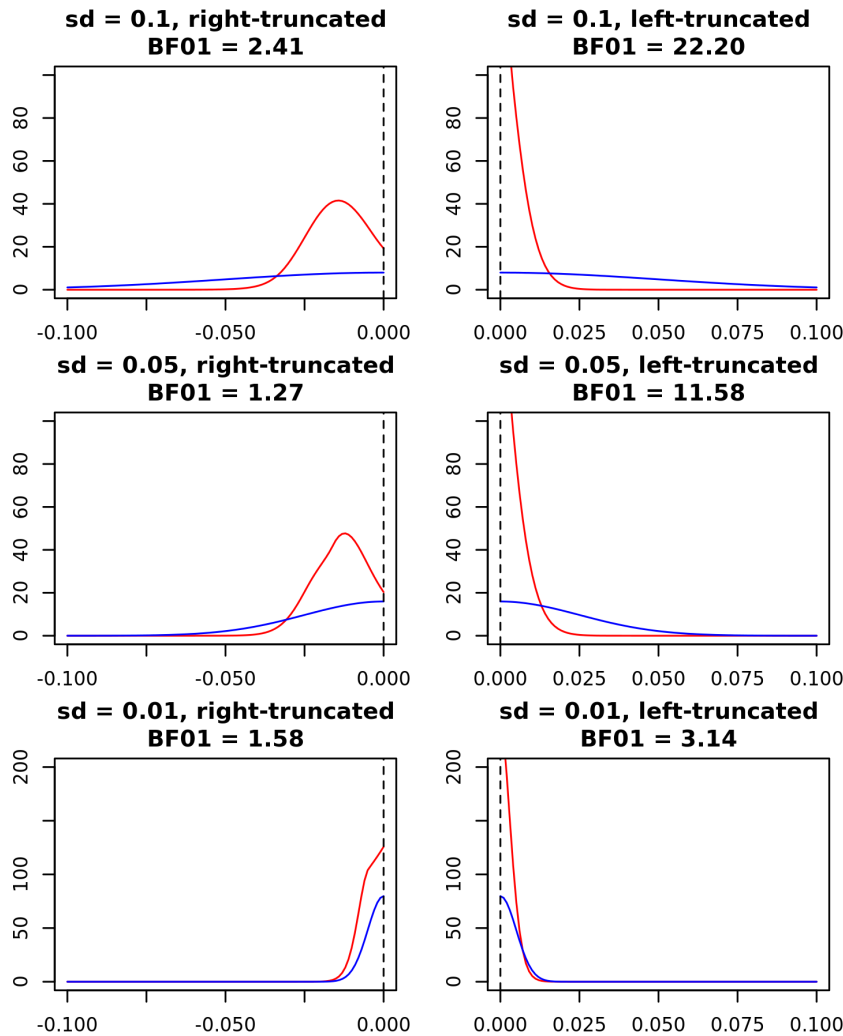
Figure 3.9: **Experiment 2 – Caterpillar plot of means and 95% credible intervals for parameters of interest, one model fit across regions crit–2 through crit+2.** *diff* parameters represent successive differences between regions, in linear order. *comp* = complexity, *p_type* = probe type.



The fact that the complexity- or length-induced speedup appeared prior to encountering the ellipsis site in Experiment 2 as well as in Experiment 1 is interesting from a methodological perspective. Remember that while the complexity manipulation introduced additional presentation regions in Experiment 1, in Experiment 2 simple and complex antecedents had the exact same number of regions. It thus seems to make no difference for the speedup effect how many presentation regions participants have passed. Rather, the quickening of the pace appears to be related to the amount of words that have been read. Keeping the number of presentation regions constant across conditions is thus not a remedy for the word-position confound that is also present in earlier studies, with the exception of Martin and McElree (2008).

The group manipulation did not appear to have any particularly strong effect on reading times for unique regions throughout the sentence, with the exception of some suggestive evidence at the region following the ellipsis. The combined analysis showed that there was a steeper drop in reading times at this position for the superficial compared to the detailed

Figure 3.10: **Experiment 2 – Density plots for the Bayes factor analysis.** Blue = prior density, red = posterior sample density.



probe group, and that afterwards reading times rose more steeply for the superficial group, returning to almost identical levels across groups. It thus appears that the detailed probe group did additional processing in this region, possibly due to spillover from the preceding ellipsis region. Indeed, the region-by-region analysis revealed suggestive evidence that the detailed probe group spent more time on region crit+1, irrespective of antecedent complexity. Speculatively, spillover might have been a factor in Experiment 2 as opposed to Experiment 1 due to the switch to centered presentation: the latter mode may increase memory demands due to the absence of visual cues (in the form of underscores) to the surrounding linguistic context. The main effect of probe type may then be due to members of the detailed probe

group allowing themselves more time to finish the antecedent-ellipsis integration, knowing that their interpretation would be queried later.

General Discussion

We have reported two studies on antecedent complexity effects in ellipsis processing. Experiment 1 yielded evidence that increasing antecedent complexity did not influence reading times at the ellipsis site, but showed that if there is such an influence, it is unlikely to be in the form of a speedup, contra Hofmeister (2011). Similarly, the results of Experiment 2 showed no effects of antecedent complexity that would have been limited to the ellipsis site, as well as no interaction between antecedent complexity and the difficulty of the end-of-sentence probe task. Given the persistent overall speedup effect that was visible in Experiment 2, we take the results of this study to be at odds with the predictions of Murphy (1985). However, the pointer model of Frazier and Clifton (2001, 2005) and Martin and McElree (2008) is able to account for both of the findings, as the proposed memory retrieval mechanism is insensitive to antecedent complexity in terms of retrieval time.

For Experiment 1, the analysis of region-by-region data revealed that while increased antecedent complexity generally led to a decrease in reading times for the rest of the sentence, this effect was suspended at the critical region, both for ellipsis and control sentences. We have suggested that the use of *and* may have caused readers to assume parallelism between the conjuncts and created an expectation of a causal connection between the first and second clauses, leading to either maintenance or retrieval and subsequent integration of material from the first conjunct at the critical region across the board.

With regard to the account of Hofmeister (2011), our current findings show that even if certain kinds of elaboration can aid retrieval, adding genitive modifiers to object noun phrases inside VP antecedents (Experiment 2), and adverbials and modal verbs to clauses (Experiment 1), do not appear to constitute cases of such ‘helpful’ elaboration. Whether this is a desirable corollary for the theory remains to be determined in future work.

In both studies, we observed an overall decrease in reading times in the regions between a longer, more complex antecedent and the ellipsis site. This pattern by itself is not new or surprising (Demberg & Keller, 2008; Ferreira & Henderson, 1993). Nevertheless, our results indicate that it does not matter in terms of the length-induced speedup if the lengthening occurs within one presentation region, as in Experiment 2, or if extra regions are added to

the sentence, as in Experiment 1, which suggests that it is words, not button presses, that make people increase their reading speed over time.

If we take the difference in parallelism requirements between the conjunctions *and* and *but* seriously, the speedup also does not seem to be related to parallelism, but may have a more mundane explanation. The working memory model of Just and Carpenter (1992) assumes that sentence comprehension involves a constant trade-off between storage and processing. A reader who has already stored more information in his or her working memory will have fewer resources available to devote to the processing of incoming words. The standard view is that reading times should increase, as it takes longer to accomplish the same task with fewer resources. Given previous work on the influence of task demands on linguistic processing, however, one may ask if the reader might not benefit from speeding up instead of slowing down. The parsing model of Lewis and Vasishth (2005), for example, assumes that linguistic information in working memory is subject to interference and decay effects which diminish the quality of the traces as new material comes in. If the participant strives to keep these traces intact, either in order to be able to answer comprehension questions or to be better able to integrate early- with late-arriving information, it may make sense to increase reading speed up to some threshold.

Conclusions

Experiment 1 yielded evidence against the assumption that increased antecedent complexity leads to faster processing of ellipsis (Hofmeister, 2011). Rather, the effect of antecedent complexity is most likely null, as predicted by pointer-based accounts of ellipsis processing (Frazier & Clifton, 2001, 2005; Martin & McElree, 2008) or otherwise a numerically very small slowdown, as would be predicted by the account of Murphy (1985). However, the results Experiment 2 call the possibility of a slowdown into question, as no such effect became visible even when task demands were high. Still, several qualifications are in order. Ellipsis is not processed in a vacuum: sentence context and discourse relations between antecedent and ellipsis clause may enhance or mask subtle effects of complexity on retrieval, and/or interact with the manipulation themselves. It might also be that different types of antecedent complexity influence retrieval times at the ellipsis site to different degrees. Murphy's (1985) assumption of a string-copying procedure would predict the length of the antecedent to be most important, while other accounts assume that the ellipsis gap contains syntactic structure (e.g., Frazier & Clifton, 2001; Merchant, 2001), which would point towards factors like the number of syntactic phrases being critical. Still other accounts may claim that as ellipsis is a discourse phenomenon, and thus makes reference to a discourse model

(e.g., Hardt, 1993), the number of unique discourse referents contained in the antecedent may play a role. In future work, we suggest to manipulate these aspects independently in order to distinguish different theories of ellipsis processing more clearly.

Acknowledgements

The authors would like to thank Brian Dillon for making it possible to run Experiment 2 at the University of Massachusetts, Amherst, as well as Johanna Thieke and Shayne Sloggett for their assistance with data collection. The insightful comments of two anonymous reviewers helped to greatly improve the article over a previous version. Experiment 1 was funded by the University of Potsdam.

CHAPTER 4

PROCESSING OF ELLIPSIS WITH GARDEN-PATH ANTECEDENTS IN FRENCH AND GERMAN: EVIDENCE FROM EYE TRACKING

Dario Paape, Barbara Hemforth, and Shravan Vasishth

SUBMITTED TO PLOS ONE

Abstract

In a self-paced reading study on German sluicing, Paape (2016) found that reading times were shorter at the ellipsis site when the antecedent was a temporarily ambiguous garden-path structure. As a post-hoc explanation of this finding, Paape assumed that the antecedent's memory representation was reactivated during syntactic reanalysis, making it easier to retrieve. In two eye-tracking experiments, we subjected the reactivation hypothesis to further empirical scrutiny. Experiment 1, carried out in French, showed no evidence in favor of the reactivation hypothesis. Instead, results for one out of the three types of garden-path sentences that were tested suggest that subjects sometimes failed to resolve the temporary ambiguity in the antecedent clause, and subsequently failed to resolve the ellipsis. The results of Experiment 2, a conceptual replication of Paape's original study carried out in German, are compatible with the reactivation hypothesis, but leave open the possibility that the observed speedup for ambiguous antecedents may be due to occasional retrievals of an incorrect structure.

Introduction

Ellipsis has received considerable attention in both theoretical linguistics and experimental psycholinguistics. Part of the appeal of ellipsis is that meaning is essentially generated from nothing. In (21), the second clause attains the meaning *but I don't know what Jane was supposed to do* without the lexical items being present. This particular type of ellipsis, in which only a wh-pronoun is left behind, is known as 'sluicing' (Ross, 1969).

(21) Jane was supposed to do something, but I don't know what.

The generation of meaning from silence is difficult to capture using traditional phrase structure rules, in which syntactic phrases are projected from a layer of terminal symbols taken from the mental lexicon. Having the word *what* project a whole clause would violate the principle of syntactic headedness, as *what* is not of the required syntactic category, not to mention that invisible terminal symbols would be needed to carry the meaning conveyed by the ellipsis (Williams, 1977). An alternative view would assume that there is an invisible pro-form at the ellipsis gap, much like a trace left behind by syntactic movement, that is co-indexed with the antecedent clause (e.g., Hardt, 1993).

From a processing perspective, the 'something from nothing' challenge becomes less of a problem: upon encountering the wh-pronoun at the end of the clause, the reader notices that there must be a gap, and the parser simply has to access the relevant material from the antecedent in order to fill in the elided part of the second clause. There is a long-standing debate about the exact nature of the 'filling in' process, which has been claimed to produce syntactic structure at the ellipsis site (Chung, Ladusaw, & McCloskey, 1995; Frazier & Clifton, 2001) or involve only a transfer of meaning (Hardt, 1993). Both accounts, however, lead to empirical problems, as noted by Merchant (2013), *inter alia*.

An issue that is sometimes framed as being part of the syntax/semantics dichotomy, but is, in fact, orthogonal to it (Phillips & Parker, 2014) concerns the mechanism by which information from the antecedent is transferred to the ellipsis site. There is converging evidence from processing studies on ellipsis that neither the size of the antecedent nor the distance between antecedent and gap influence the time it takes to resolve the ellipsis (Frazier & Clifton, 2000, 2001; Martin & McElree, 2008, 2009) (see, however, Murphy, 1985 for a diverging result). The observed pattern suggests that the parser has direct access to the antecedent's representation in memory, without having to initiate a time-consuming search process, and that all information from the antecedent is accessed at once, so that processing is not slowed for bigger as compared to smaller antecedents.

A recent study by Paape (2016) investigated whether ellipsis is processed differently depending on whether the antecedent contains a temporary syntactic ambiguity or not. A German example stimulus from the study, adapted slightly in favor of brevity, is shown in (22).

- (22) Eine Sympathisantin der Opposition hatten die Rebellen unterstützt,
A sympathizer.fem.nom/acc of the opposition had.pl the rebels supported,
aber niemand weiß, wie, ...
but nobody knows how
‘The rebels had supported a sympathizer of the opposition, but nobody knows how, ...’

Here, the first noun phrase *Eine Sympathisantin ...*, ‘a sympathizer ...’, is initially ambiguous between nominative and accusative marking. German readers prefer to adopt a subject reading for this noun phrase (Hemforth, 1993; Meng & Bader, 2000) and are thus forced to reanalyze when the plural-marked auxiliary *hatten*, ‘had.pl’, arrives to indicate that the sentence has non-canonical OVS order. In a self-paced reading experiment, Paape (2016) found that reading times were increased at the second noun phrase in (22) compared to a control sentence in which the auxiliary bore singular agreement, which disambiguates the clause towards the preferred reading. No such effect was observed in two control sentences where the first noun phrase carried overt case marking, indicating that readers experienced a garden path upon disambiguation in the ambiguous OVS condition.

The use of potential garden-path sentences as antecedents for ellipsis was intended as a more stringent test of the assumption that the parser has direct access to a structured representation of the antecedent. If reanalysis of the ambiguous structure is successful – a reasonable assumption given that OVS constructions are not particularly rare in German – the parser should have a fully specified antecedent structure available when the ellipsis site is encountered. As this should also be the case in the control conditions, no difference in processing times is predicted. Alternatively, the parser may only have access to a (partly) unstructured antecedent representation, that is, a loose collection of constituents or words, and would have to compute the necessary syntactic (or semantic) structure *again* at the ellipsis site, which would possibly result in a reappearance of the garden-path effect.

Apart from replicating the classic garden-path effect in the antecedent clause, Paape’s (2016) study yielded two main results: sentences with non-canonical OVS order in the antecedent showed longer reading times in a spillover region two regions downstream from the ellipsis site, independently of ambiguity. Three regions downstream from the ellipsis site, there was an interaction between case marking on the initial noun phrase and antecedent word order, such that sentences with ambiguous OVS antecedents were processed faster compared to

both their ambiguous SVO counterparts and control sentences. Furthermore, sentences with unambiguous OVS antecedents showed the longest reading times at this position.

Paape's findings are incompatible with the view that the structure of the antecedent has to be recomputed at the ellipsis site: such an account would have predicted the opposite pattern, namely increased reading times for ambiguous OVS sentences at the ellipsis site, given that the garden-path effect should have been experienced again. The results are, however, in line with the retrieval of a stored structure if additional assumptions are made.

In the cue-based retrieval model of parsing developed by Lewis and Vasishth (2005), syntactic chunks with varying activation levels are retrieved from memory in order to form new structures. Each chunk is assumed to have an associated activation level that will decrease with time unless the chunk is retrieved, that is, unless it is needed for syntactic computation. Any chunk that *is* retrieved receives a boost to its activation level. Even though this boost decays over time, it is permanent in the sense that the activation of a chunk that has been retrieved n times will not decay below the activation level of a chunk that has been retrieved $n - 1$ times.

Under the Lewis and Vasishth model, a sluicing gap would set a retrieval probe for a fully parsed clause in memory. The latency of the retrieval decreases with the target's activation, which is determined by its retrieval history (see above) and the match between the target's feature specification and the features cued for by the probe. Paape (2016) suggests that the reanalysis of ambiguous OVS antecedents indexed by the garden-path effect in the antecedent clause involves additional retrievals of the corresponding memory chunk. This results in an activation boost, leading to faster retrieval at the ellipsis site later on. As no reanalysis and therefore no reactivation is assumed for ambiguous SVO or unambiguous antecedents, their activation levels are predicted to be lower upon retrieval, resulting in slowed processing compared to ambiguous OVS antecedents.

However, OVS antecedents also appeared to slow down processing at the ellipsis site, as indicated by the increased reading times at the second spillover region and the high reading times for sentences with unambiguous OVS antecedents at the third spillover region. In order to explain the finding, Paape (2016) notes that there is a form mismatch between antecedent and ellipsis site in (22): the wh-pronoun *wie*, 'how', marks the beginning of a subordinate clause – the complement of *weiß*, 'knows' – which in German would canonically show SOV word order if there was no ellipsis, see (23) below. Note that an unelided utterance would use a definite rather than an indefinite article, and most likely drop the modifier.

- (23) [Eine Sympathisantin der Opposition]_S hatten_V [die Rebellen]_O
A sympathizer.fem.nom/acc of the opposition had.pl the rebels
unterstützt, aber niemand weiß, wie [die Rebellen]_S [die Sympathisantin]_O
supported, but nobody knows how the rebels the sympathizer
unterstützt hatten_V.
supported had
‘The rebels had supported a sympathizer of the opposition, but nobody knows how
the rebels had supported the sympathizer.’

Given that the antecedents in Paape’s study had either SVO or OVS word order, there was thus a mismatch between the feature set requested by the ellipsis gap, which would by assumption probe for an SOV feature, and the antecedent’s feature specification. Nevertheless, Paape argues that an SVO antecedent provides a better match for retrieval cues set by the ellipsis than an OVS antecedent, given that the order of subject and object matches the gap’s specifications.

One caveat is that OSV is actually a permitted, though dispreferred, word order in German subordinate clauses. As the linear order mismatch between OSV and OVS is the same as between SOV and SVO, assuming that the gap probes for any permitted word order would predict no disadvantage for OVS antecedents. In order to derive the correct predictions, one must thus subscribe to the assumption that only the canonical word order serves as a retrieval cue.

As Paape’s (2016) conclusions were arrived at post-hoc, it is important to subject them to further empirical investigation. The critical interaction between antecedent ambiguity and word order was found only in a late spillover region and after analyzing multiple regions of interest, which leaves open the possibility that the result is a false positive (type I error) (von der Malsburg & Angele, 2017). Moreover, the observed power of Paape’s experiment – which is likely an overestimate of the actual power (Vasisht & Gelman, 2017) – was at 45%, which adds the possibility of a false negative result (type II error) with regard to potential effects at the ellipsis site.

Below, we present two new studies which were designed to test the reactivation hypothesis. Instead of self-paced reading, both of our experiments used eye tracking during reading, which arguably provides a more naturalistic way of presenting the stimuli. Self-paced reading is known to frequently show spillover effects, suggesting that participants may choose to hold words in memory rather than integrate them immediately. Such a strategy may be related to the adoption of a fixed rhythm in pressing the space bar (Koornneef & Van Berkum, 2006; Witzel et al., 2012), which could obscure processing patterns that may become visible when the stimulus presentation more closely resembles normal reading.

Importantly, the eye-tracking paradigm also provides subjects with the opportunity to make regressions to earlier parts of the sentence. One concern in relation to the original result was that the speedup observed for ambiguous OVS antecedents could be due to parsing failure rather than easier retrieval: if subjects fail to resolve the garden path and thus do not create a well-formed antecedent memory chunk, they may subsequently fail to resolve the ellipsis, given that they have no chance to return to the antecedent clause and reread it (though making a ‘covert’ regression is a logical possibility). The original study yielded no direct evidence for such failures: while comprehension accuracy suffered in the garden-path condition, questions targeting the interpretation of the antecedent and questions targeting other parts of the stimulus sentence were affected to the same degree, suggesting that parsing failure may not have been responsible for the errors. In eye tracking, targeted regressions to the antecedent in ambiguous sentences would indicate that the unavailability of a memory target may indeed play a role in creating the observed speedup at the ellipsis site, and that readers are making a second attempt at reanalyzing the antecedent.

Experiment 1

The design of our first study differs from Paape’s (2016) original experiment in three respects:

- I) Language: Our experiment was carried out in French rather than German. Concordant evidence from the two languages would strengthen the claim that reactivation through reanalysis is a language-general phenomenon.
- II) Method: We used eye tracking instead of self-paced reading.
- III) Type of syntactic ambiguity: In addition to sentences with non-canonical word order which closely resembled Paape’s original German stimuli, our design also incorporated sentences containing main clause/reduced relative and lexical ambiguities, as explained below.

If the reactivation hypothesis is correct, the observed pattern of faster processing of garden-path antecedents at the ellipsis site should be reproducible even when these parameters are changed.

Materials

Like German, French allows sluicing, as shown in (24).

- (24) Je veux acheter un cadeau, mais je ne sais pas lequel.
I want to buy a present but I NEG know NEG which (one)
'I want to buy a present, but I don't know which one.'

By investigating possible antecedent reactivation effects in French sluicing constructions, we are thus looking at the same syntactic phenomenon investigated by Paape (2016) in a different linguistic environment.

Given that syntactic reanalysis is critical to the reactivation hypothesis, the first challenge was to identify suitable French garden-path sentences, which are comparatively understudied. The few studies we could find investigated relative clause attachment ambiguities (Zagar, Pynte, & Rativeau IV, 1997), ambiguous wh-questions with long-distance extraction (Lassotta, Omaki, & Franck, 2016) and a particular type of causative/non-causative ambiguity which is limited to the verb *faire*, 'do' (Frenck-Mestre, 2002). None of these seemed suitable for our needs in the current study. Specifically, relative clause attachment has been argued to possibly involve underspecification (Swets et al., 2008), the grammatical status of long-distance extractions is dubious with some verb types (Ambridge & Goldberg, 2008), and the ambiguity of *faire* was deemed too lexically specific.

We identified three potential garden-path constructions in French which to our knowledge have not been investigated systematically before: sentences featuring subject-object inversion (SOI stimuli), sentences containing reduced relative clauses (RRC stimuli) and sentences containing a sequence of three lexically ambiguous words (triple lexical ambiguity, TLA stimuli). Apart from testing the reactivation hypothesis, the current study had the additional aim of establishing whether garden-path effects do indeed occur with these constructions. Each of the three sentence types is discussed separately below.

The 56 experimental stimuli ($20 \times \text{SOI} + 20 \times \text{RRC} + 16 \times \text{TLA}$) were combined with 24 filler sentences from an unrelated experiment, yielding a total of 80 sentences. Items were rotated through the experimental conditions according to a latin-square design, so that no subject saw the same sentence twice.

Type 1: Subject-object inversion (SOI)

Stimuli of the first type were wh-questions involving simple main clauses in which the order of subject and object was either canonical (SVO) or reversed (OVS), as shown in (25). Diamonds indicate region-of-interest boundaries in all examples. The point of disambiguation and the critical ellipsis regions are marked by subscripts.

- (25) Quel escrimeur ◊ de l'équipe nationale ◊ a / ont salué_{disambig.} ◊ les
Which fencer of the team national has / have greeted the
adversaires ◊ du concours ◊ et quand / avant la lutte_{critical} ◊ si je puis
opponents of the contest and when / before the fight if I can
me permettre?
SELF allow
'Which fencer of the national team greeted the opponents in the contest / did
the opponents in the contest greet and when / before the fight, if I may ask?'

This sentence type almost directly parallels the German construction used by Paape (2016). The first noun phrase (*Quel escrimeur ...*, 'which fencer ...') could initially be either a subject or an object, though the subject reading is more common, and thus likely to be preferred by French readers, as it is by readers of other languages Bever (1970); Townsend and Bever (2001).

One should note that claiming SVO to be *the* canonical word order in French is clearly problematic when sentences with clitics are considered Charvillat and Kail (1991), but significantly less so when all argument roles are filled by lexical noun phrases, as in the present study.

In (25), when the auxiliary bears singular marking (*a*, 'has'), it agrees with the initial noun phrase, which is thus identified as the clause's subject. When it is plural (*ont*, 'have') it agrees with the plural-marked second noun phrase (*les adversaires ...*, 'the opponents ...'), which indicates that the initial noun phrase must be the direct object of *salué*, 'greeted'.

The initial clause is either coordinated with a sluice (*quand*, 'when') via the conjunction *et*, 'and', in the ellipsis conditions. In the control conditions, it continues with an adverbial phrase (*avant la lutte*, 'before the fight'). Items of this type thus followed a 2 × 2 design with the factors canonicity (canonical vs. non-canonical) and elision (ellipsis vs. control). Each item ended with a set phrase (e.g., '... if I may ask') which was intended as a spillover region.

A total of 20 SOI sentences were used in the experiment. In half of these the first noun phrase was singular and the second noun phrase was plural, as in (25), while for the other

half the gender features were switched, making plural marking on the auxiliary the form that agreed with the initial subject.

Type 2: Reduced relative clauses (RRC)

For the second stimulus type, we made use of the fact that some French verbs show syncretism between their third-person active and participle forms, such as *détruit*, ‘destroyed’, in (26). As in comparable English sentences, whose processing has been studied quite extensively (e.g., Ferreira & Clifton, 1986; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993), this leads to a temporary ambiguity between an active reading of the initial clause (*The boat destroyed [something]*) and a reading in which *détruit* heads a reduced relative clause (RRC) modifying the initial noun. The sentence is disambiguated towards the latter reading at *avait rejoint*, ‘had returned (to)’, which marks the end of the relative clause. In French, an adverb can intervene between an active verb and its direct object, so the prepositional phrase *pendant la guerre*, ‘during the war’, does not disambiguate the sentence.

(26) Ambiguous antecedent

Le navire ◊ détruit ◊ pendant la guerre ◊ avait rejoint_{disamb.} ◊ le port,
 The ship destroyed during the war had returned.to the harbor

Unambiguous antecedent

Les navires
 The.PL ships
 ◊ détruits ◊ pendant la guerre ◊ avaient rejoint_{disambig.} ◊ le port,
 destroyed.PL during the war had.PL returned.to the harbor

‘The ship(s) (that was/were) destroyed during the war had returned to the harbor, ...’

Ellipsis

... mais le professeur ◊ d’histoire ◊ ne pouvait pas dire ◊ quand_{critical}
 but the professor of history NEG could NEG say when

Control

... mais le professeur ◊ d’histoire ◊ n’en savait rien_{critical}
 but the professor of history NEG.of it knew nothing

‘... but the history professor could not say when / knew nothing about it ...’

... laissant
leaving
◇ un peu désillusionnée ◇ la jeune collègue ◇ durant leur rendez-vous.
a bit disappointed the young colleague during their date
'... leaving the young colleague a bit disappointed during their date.'

The unambiguous version of the sentence is obtained by explicitly marking the participle status of the verb form through number or gender agreement with the initial noun phrase. Note that while an active verb in this position would also agree with initial noun phrase, a different morphological paradigm would apply, and the corresponding form in (26) would be *détruisent*. Besides the ambiguity factor (ambiguous vs. unambiguous), RRC stimuli also featured an elision manipulation (ellipsis vs. control), again resulting in a 2×2 design.

In the ellipsis conditions, the clause headed by *mais*, 'but', ends with a sluice (... *quand*, 'when'). The control conditions for this sentence type always contained a non-elliptical anaphor, such as *en*, 'of.it', in the above example. Across items, the anaphor sometimes appeared in the same region as the ellipsis, but could also appear at an earlier point (see item list in Appendix IV). The rest of the sentence served as a spillover region, and to keep subjects interested by providing a more elaborate discourse. A total of 16 RRC items were used in the experiment.

Type 3: Triple lexical ambiguity (TLA)

Items of the third stimulus type contained sequences of three lexically ambiguous words. In example (27), the translation of the first clause is 'The filthy butcher cuts them'. The pronoun *les*, 'them', refers to the big sides of beef mentioned in the context sentence which accompanied the item.

(27) **Context**

Le marché du quartier
The market of the district
est connu pour les grandes côtes de bœuf qui sont livrées durant la nuit.
is known for the big sides of beef which are delivered during the night

Ambiguous antecedent

Le boucher ◇ sale ◇ les tranche_{,disambig.}
the butcher filthy them cuts

Unambiguous antecedent

Les bouchers ◊ sales ◊ les tranchent_{,disambig.}
 the.PL butchers filthy.PL them cut

‘The filthy butcher(s) cut(s) them ...’

Ellipsis

... mais les clients
 but the clients
 ◊ de la place Colbert ◊ se demandent ◊ incroyablement ◊ quand_{,critical}
 of the square Colbert SELF ask disbelievingly when

Control

... mais les clients ◊
 but the clients
 de la place Colbert ◊ en demandent ◊ incroyablement ◊ la technique_{,critical}
 of the square Colbert of.it ask disbelievingly the technique

‘... but the clients of Colbert square ask disbelievingly when / about the technique of it, ...’

vu que la viande
 seen that the meat
 est vendue déjà marinée et prête à l’emploi en début de matinée.
 is sold already marinated and ready for the.use at beginning of morning
 ‘... given that the meat is sold already marinated and ready to use in the early morning.’

The words *sale*, *les* and *tranche* are all lexically ambiguous: *sale* is both the singular form of the adjective ‘filthy’ and the third-person singular form of the verb *saler*, ‘to salt’, while *les* can be either a plural-marked direct object pronoun or a plural-marked definite article. Besides being the third-person singular form of the verb ‘to cut’, *tranche* can also be a noun (‘slice’).

The alternative reading resulting from this chain of ambiguities, namely ‘The butcher salts the slices’, which is also compatible with the context sentence, is shown in (28). Note, however, that the would-be noun *tranche* would need to bear a plural suffix (-s) in order for this alternative reading to be grammatical. The stimulus sentence shown above in (27) is thus disambiguated towards the other reading by the absence of the plural marker.

(28) **Garden-path reading of (27)**

Le boucher ◊ sale ◊ les tranches, ...
the butcher salts the slices

TLA stimuli also followed a 2×2 design with the factors ambiguity (ambiguous vs. unambiguous) and elision (ellipsis vs. control). The ambiguous reading becomes impossible when the initial noun phrase is plural and the adjective *sale*, ‘filthy’, bears a plural -s. Again, the third-person plural form of the verb *sale*, ‘to salt’, which is *salent*, would look different from the plural form of the adjective. The verb form *tranchent*, ‘cut’, also agrees with the plural feature of the initial noun phrase in the unambiguous conditions.

As for the RRC stimuli, the second clause, introduced by *mais*, ‘but’, contained either a sluice (... *quand*, ‘when’) or a pronoun (*en*, ‘of it’). Again, anaphor and ellipsis sometimes appeared within the same region for a given item, but this was not always the case. As before, the remainder of the sentence was intended both as a spillover region and as a reasonably interesting continuation of the discourse. The experiment contained 20 TLA sentences, each accompanied by a context sentence.

Predictions

A vital assumption of the current design is that syntactic reanalysis of the antecedent clause is required across all three stimulus types. For SOI sentences, subjects are expected to initially compute an SVO structure for all sentences, but the non-canonical versions force a switch to OVS at the auxiliary, much like in the German sentences of Paape (2016). For ambiguous TLA and RRC sentences, we assume that readers will initially adopt the preferred structure, which in both cases corresponds to an SVO active sentence, and are forced to abandon this analysis once the disambiguating word – either the singular-marked noun in TLA sentences or the finite verb in RRC sentences – arrives. We thus expect processing difficulty, in the form of elevated reading times that may be accompanied by regressions, at the disambiguating region across all stimulus types.

If the reactivation hypothesis is correct, reanalyzed antecedents should be easier to retrieve at the sluicing site in the elided conditions, leading to shorter reading times and possibly fewer and/or shorter regressions. In contrast, we do not expect to see an advantage for ambiguous/non-canonical antecedents in the control conditions, as pronouns are assumed to be resolved by discourse-based mechanisms instead of by the retrieval and integration of syntactic chunks from working memory (Hardt, 1993).

Pre-study: Off-line antecedent acceptability ratings

In order to assess whether garden-path effects were likely to be observed for our three stimulus types, we ran an Internet-based pre-study with 52 French native speakers who did not participate in the main study. The experiment was hosted on the Ixet farm (Drummond, 2017). Informed consent was obtained from each participant prior to the experiment (in written form in the case of Experiments 1 and 2). The pre-study as well as the experiments described below complied with the June 1964 Declaration of Helsinki, as last revised. At the time of experimentation, both French and German law allowed non-invasive experiments with human subjects to be conducted without prior approval from an internal review board (or ethics committee). Instead, compliance with the Declaration of Helsinki was ensured by the primary investigator at the site of each experiment, in this case Prof. Barbara Hemforth for the pre-study and Experiment 1 and Prof. Shravan Vasishth for Experiment 2.

All stimulus sentences were truncated at the first comma, such that only the antecedent clause remained, which was presented at once in its entirety. There was thus only one manipulation for each sentence type: RRC and TLA stimuli were either ambiguous or unambiguous while SOI stimuli were either canonical or non-canonical. In the case of the TLA stimuli, the context sentence was visible on the screen, but participants were explicitly instructed to evaluate only the target sentence. Stimuli were presented in random order according to the design described above and participants were required to judge the acceptability of each sentence on a scale from 1 to 10. We opted for the 1-10 scale, as opposed to the more commonly used 1-7 scale, because French subjects are used to it from their time in school.

There was no time limit for giving the rating. We expected our participants to give lower ratings to sentences which cause processing difficulty, that is, to garden-path sentences as opposed to unambiguous controls.

Data analysis

The data were analyzed using the statistics software **R** (R Core Team, 2016) and the **Stan** programming language for Bayesian statistical inference (Stan Development Team, 2016). As ratings on a Likert scale are discrete rather than continuous, and tend to show distributions that deviate from normality, we fitted a cumulative logit model across sentence types. Models of this type assume a latent variable η and a set of cutpoints C which demarcate the boundaries of each rating's 'bin' in logit space. For instance, when η assumes a value between the cutpoints for rating 1 and rating 2, a rating of 1 will be assigned. One

advantage of this approach is that the differences between any two adjacent cutpoints do not need to be of the same size, such that some manipulations may, in principle, only affect ratings within a certain range on the Likert scale.

We assumed a common set of cutpoints across sentence types as well as across items and participants. In addition to a separate intercept on η and separate coefficients for the two experimental factors by sentence type, the model included random intercept and slope adjustments both by participants and by items. The ambiguity/canonicity factor was sum-coded, such that ambiguous/non-canonical sentences were coded as 1 and unambiguous/canonical sentences as -1 . We set uninformative Cauchy(0,2.5) priors for all coefficients and an LKJ prior (Lewandowski et al., 2009) with parameter 2 for the variance-covariance matrices of the random effects for subjects and items.

Four chains with a total of 4000 iterations each and a warm-up period of 2000 iterations were run for each model. Convergence was verified based on **Stan**'s \hat{R} statistic. We only report effects for which zero is not included in the associated 95% credible interval or for which there is nevertheless a high probability ($> 95\%$) that the parameter is above or below zero.

Results

Mean ratings by sentence type and condition are shown in Table 4.1. There was no evidence of an effect of ambiguity on ratings for RRC sentences. For TLA sentences, ambiguous versions received lower ratings than unambiguous versions ($\hat{\beta} = -0.12$, CrI: $[-0.23, -0.01]$, $\Pr(\hat{\beta} > 0) = 0.02$). For SOI sentences, non-canonical versions also received lower ratings than canonical versions ($\hat{\beta} = -1.02$, CrI: $[-1.15, -0.89]$, $\Pr(\hat{\beta} > 0) \approx 0$).

Table 4.1: **Pre-study: Mean acceptability ratings and standard errors (in parantheses) by sentence type and condition.**

SOI stimuli	canonical	non-canonical
	8.46 (0.10)	5.68 (0.15)
RRC stimuli	unambiguous	ambiguous
	7.94 (0.13)	7.83 (0.13)
TLA stimuli	unambiguous	ambiguous
	6.01 (0.14)	5.75 (0.14)

Given these preliminary results, we conclude that ambiguous TLA and non-canonical SOI stimuli caused greater processing difficulty for participants, leading to lower ratings, as

intended. While there was no evidence of an effect for the RRC stimuli, this sentence type was carried over into the main experiment as well, given that a relatively mild garden path may still be detectable in an on-line experiment.

Participants

We recruited 46 French native speakers between the ages of 19 and 38 through a mailing list. All had normal or corrected-to-normal eyesight. They were paid 7 € for their participation. Informed consent was obtained from each participant prior to the experiment.

Procedure

Participants were tested individually in a dimly lit, soundproof chamber. Subjects read the experimental sentences and fillers at their own pace while an SR Research Eyelink II tracker with a head-mounted camera setup recorded the movements of their dominant eye at a sampling rate of 500 Hz. Participants were seated at a distance of 75 cm from the presentation screen.

At the beginning of each experimental session the eye-tracker was calibrated using a nine-point grid. Each trial started with the presentation of a dot on the monitor which participants had to fixate in order to start the presentation of the stimulus sentence, whose first character then appeared in the same location as the dot. Participants signaled completion of a trial by pressing a button on a provided game pad.

Sentences were presented in 16 pt Times New Roman font at a resolution of 1024×768 pixels. At a screen width of 47.5 cm, one character accounted for 0.57° of visual angle. Line breaks were coded manually for longer sentences in order to avoid lines breaking directly before or after a region of interest. A chin rest was used to keep participants' head position stable during tracking. Each session began with the presentation of two practice items to familiarize subjects with the procedure. The presentation order of the remaining sentences was randomized at runtime. All experimental items, but none of the filler items were followed by a yes/no comprehension probe. The probes targeted various parts of the sentence, including but not limited to the interpretation of the antecedent and the ellipsis. There were three obligatory breaks during the experiment, and subjects were told that they could take additional breaks at any time. The eye-tracker was recalibrated using a nine-point grid after each break. Recording sessions lasted 45 minutes on average.

Data Analysis

Due to repeated tracking loss and a computer crash, no analyzable data were collected for five participants. For the remaining 41 participants, data points in which a region had a total reading time of less than 20 ms were removed prior to the statistical analysis. Visual inspection suggested non-normality of residuals across all reported eye-tracking measures. We thus applied the Box-Cox procedure (Box & Cox, 1964) to identify an appropriate transform. The procedure suggested a logarithmic transformation of the response, hence all further analyses were carried out assuming that the dependent measures were log-normally distributed.

Our Stan model included a global standard deviation $\hat{\sigma}$ and separate intercepts $\hat{\alpha}_{sr}$ per region of interest r of sentence type s , as shown in Equation 4.1 for first-pass reading times.

$$\begin{aligned}
 \hat{\mu}_{srij} &= \hat{\alpha}_{sr} + u_i + w_{sj} + u_{sri} + w_{srj} + \\
 &\quad (\hat{\beta}_1 + u'_i) \cdot \text{length} + \\
 (4.1) \quad &\quad (\hat{\beta}_{2sr} + u''_{sri} + w'_{srj}) \cdot \text{elision} + \\
 &\quad \dots \\
 FPRT_{srij} &\sim \text{LogNormal}(\hat{\mu}_{srij}, \hat{\sigma})
 \end{aligned}$$

The intercept adjustment u_i by subject i is global in the sense that it is applied to all regions of interest across all sentence types, thus taking into account that all items were read by the same subjects. The intercept adjustment w_{sj} is applied across regions of interest for each item within each sentence type. Two further adjustments to the intercept, u_{sri} and w_{srj} , are applied by region within each sentence type, for each item j of a given type and for each subject i . Together, these adjustments take into account possible variability in reading times between subjects and items for a particular region of interest for a particular sentence type. For all analyses of eye-tracking measures, the log-transformed number of characters in each region of interest was added to the model as a centered covariate to control for length differences between conditions. The effect of length was estimated globally across items, regions and sentence types, but was allowed to vary between subjects by assuming the slope adjustment u'_i . Lastly, slope adjustments u''_{sri} and w'_{srj} for each experimental factor account for variability in the effect of the experimental manipulations between items and subjects. Like the associated coefficient $\hat{\beta}$, each adjustment is specific to region of interest r of sentence type s .

Analogous analyses, without the multiple regions of interest, were carried out for question response accuracies, which were analyzed using logistic regression, as well as for log question

response times. The accuracy model included a globally estimated coefficient with random slopes by subject for log response time to account for the interdependence between accuracy and response latency.

Across the three stimulus types, elision conditions were coded as 1 and control conditions as -1 . For RRC and TLA stimuli, the ambiguous conditions were coded as 1 and the unambiguous conditions as -1 . For SOI stimuli, non-canonical word order was coded as 1 and canonical word order as -1 . Due to experimenter error, no data on question response accuracy and response time was recorded for the first three participants, thus the reported results refer to data from the 38 remaining participants.

For nested comparisons as well as for the exploratory analyses, we used the `rstanarm` package (Gabry & Goodrich, 2016), which interfaces R with Stan and simplifies model specification by emulating the syntax of the more commonly used `lme4` package (Bates, Mächler, et al., 2015).

As before, we set $\text{Cauchy}(0, 2.5)$ priors for all coefficients. An LKJ prior with parameter 2 was used for the variance-covariance matrices of the random effects for subjects and items. Four chains with 4000 iterations each were run for each model.

Results – Subject-object inversion (SOI)

Length-corrected reading time measures by region of interest for the SOI stimuli are shown in Figure 4.1.

Comprehension probes

Table 4.2 shows the results for comprehension accuracy and response latency. The mean response accuracy for SOI stimuli was 74%. There were no effects of the experimental manipulations on response times. There was, however, an effect of the canonicity manipulation on response accuracy, such that less accurate responses were given after non-canonical sentences compared to canonical sentences ($\hat{\beta} = -0.95$, CrI: $[-1.53, -0.42]$, $\text{Pr}(\hat{\beta} > 0) \approx 0$). There was also an effect of elision on response accuracy, such that more accurate responses were given after elided sentences ($\hat{\beta} = 0.32$, CrI: $[0.01, 0.64]$, $\text{Pr}(\hat{\beta} > 0) = 0.98$).

Figure 4.1: **Reading measures by region for SOI stimuli (Experiment 1).** All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.

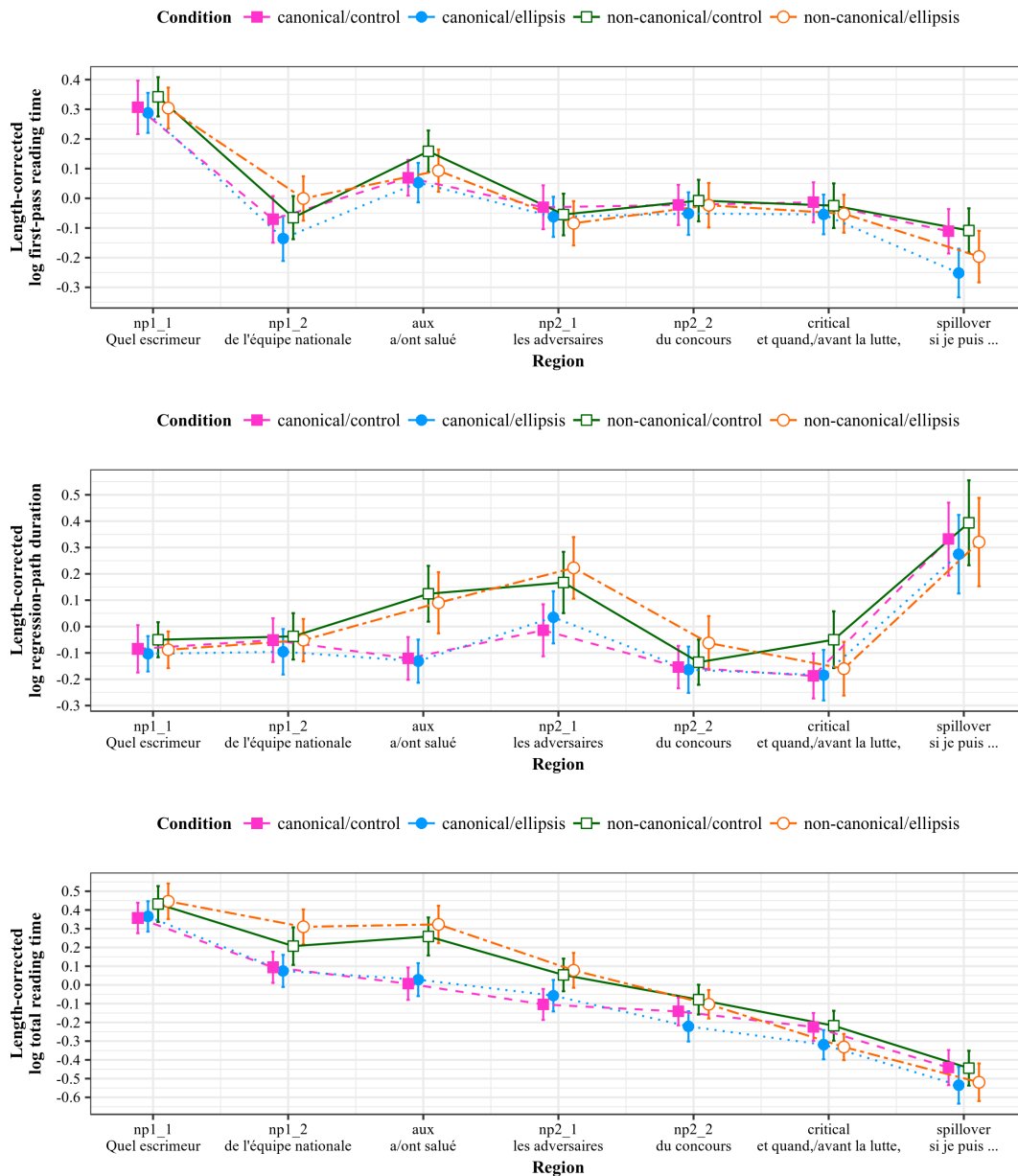


Table 4.2: **Experiment 1: Results for question responses accuracy and response times (SOI stimuli).** can = canonicity, el = elision.

response accuracy (logit scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	0.32	0.01	0.64	0.98
can	-0.95	-1.53	-0.42	≈ 0.00
el \times can	-0.12	-0.46	0.22	0.25
response time (log scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	0.00	-0.05	0.04	0.45
can	0.03	-0.03	0.09	0.82
el \times can	-0.02	-0.07	0.03	0.19

First-pass reading times

Results for first-pass reading times are shown in Table 4.3. There was an unexpected interaction between elision and canonicity in the second region of the first noun phrase ($\hat{\beta} = 0.03$, CrI: [0.00,0.07], $\Pr(\hat{\beta} > 0) = 0.96$). This effect is likely to be spurious, given that elided and unelided sentences did not differ at this point and there was no parafoveal preview of the critical region. First-pass reading times were shorter for the elided than for the non-elided conditions in the spillover region following the critical region ($\hat{\beta} = -0.05$, CrI: [-0.10,0.00], $\Pr(\hat{\beta} > 0) = 0.02$).

Regression-path durations

Table 4.4 lists the results for regression-path durations by region. The region containing the disambiguating auxiliary and the participle showed an effect of canonicity, such that non-canonical sentences showed longer regression-path durations ($\hat{\beta} = 0.11$, CrI: [0.06,0.17], $\Pr(\hat{\beta} > 0) \approx 1$). This effect also appeared in the first region of the second noun phrase ($\hat{\beta} = 0.09$, CrI: [0.03,0.16], $\Pr(\hat{\beta} > 0) \approx 1$).

Total reading times

Table 4.5 shows the results for total reading times. Total reading times were elevated in non-canonical as compared to canonical sentences for the first ($\hat{\beta} = 0.04$, CrI: [0.00,0.08],

Table 4.3: **Experiment 1: Results for first-pass reading times (SOI stimuli).** can = canonicity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1_1 <i>Quel escrimeur</i>	el	-0.01	-0.06	0.03	0.27
	can	0.01	-0.03	0.05	0.71
	el × can	-0.01	-0.05	0.04	0.38
np1_2 <i>de l'équipe nationale</i>	el	0.00	-0.05	0.04	0.45
	can	0.03	-0.01	0.08	0.92
	el × can	0.03	0.00	0.07	0.96
aux <i>a/ont salué</i>	el	-0.02	-0.06	0.02	0.17
	can	0.03	-0.01	0.07	0.96
	el × can	-0.01	-0.05	0.03	0.26
np2_1 <i>les adversaires</i>	el	-0.02	-0.07	0.04	0.28
	can	-0.01	-0.05	0.02	0.25
	el × can	0.00	-0.04	0.04	0.48
np2_2 <i>du concours</i>	el	-0.01	-0.05	0.03	0.28
	can	0.01	-0.04	0.06	0.65
	el × can	0.00	-0.04	0.04	0.45
crit <i>et quand,/avant la lutte,</i>	el	-0.02	-0.08	0.04	0.21
	can	0.00	-0.05	0.04	0.42
	el × can	0.01	-0.03	0.04	0.61
spillover <i>si je puis ...</i>	el	-0.05	-0.10	0.00	0.02
	can	0.01	-0.04	0.06	0.65
	el × can	0.01	-0.03	0.06	0.70

$\Pr(\hat{\beta} > 0) = 0.95$) and second regions of the initial noun phrase ($\hat{\beta} = 0.09$, CrI: [0.04,0.13], $\Pr(\hat{\beta} > 0) \approx 1$). The effect continued into the auxiliary region ($\hat{\beta} = 0.14$, CrI: [0.08,0.20], $\Pr(\hat{\beta} > 0) \approx 1$) as well as the first ($\hat{\beta} = 0.07$, CrI: [0.03,0.12], $\Pr(\hat{\beta} > 0) \approx 1$) and second regions of the second noun phrase ($\hat{\beta} = 0.04$, CrI: [0.00,0.09], $\Pr(\hat{\beta} > 0) = 0.98$).

Results – Reduced relative clauses (RRC)

Length-corrected reading time measures by region of interest for the RRC stimuli are shown in Figure 4.2.

Table 4.4: **Experiment 1: Results for regression-path durations (SOI stimuli)**. can = canonicity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1_1 <i>Quel escrimeur</i>		See first-pass reading times			
np1_2 <i>de l'équipe nationale</i>	el	-0.02	-0.06	0.03	0.25
	can	0.02	-0.03	0.07	0.77
	el × can	0.01	-0.04	0.06	0.63
aux <i>a/ont salué</i>	el	-0.01	-0.06	0.04	0.38
	can	0.11	0.06	0.17	≈1.00
	el × can	-0.01	-0.06	0.04	0.36
np2_1 <i>les adversaires</i>	el	0.03	-0.04	0.10	0.78
	can	0.09	0.03	0.16	≈1.00
	el × can	0.00	-0.05	0.05	0.52
np2_2 <i>du concours</i>	el	0.01	-0.04	0.07	0.71
	can	0.03	-0.03	0.08	0.86
	el × can	0.02	-0.04	0.07	0.73
crit <i>et quand,/avant la lutte,</i>	el	-0.04	-0.12	0.05	0.20
	can	0.04	-0.03	0.11	0.87
	el × can	-0.02	-0.09	0.04	0.26
spillover <i>si je puis ...</i>	el	-0.03	-0.13	0.06	0.23
	can	0.02	-0.06	0.10	0.70
	el × can	-0.01	-0.12	0.10	0.43

Comprehension probes

Results with regard to comprehension accuracy and response time are shown in Table 4.6. The mean probe response accuracy for the RRC stimuli was 86%. There were no effects of the experimental manipulations on response times. Response accuracy showed an effect of elision, such that questions about elided sentences were answered more accurately ($\hat{\beta}=0.29$, CrI: $[-0.05,0.66]$, $\Pr(\hat{\beta} > 0)=0.95$).

First-pass reading times

Table 4.7 lists the results for first-pass reading times. In the post-disambiguating region, there was an effect of ambiguity, such that ambiguous sentences were read faster than unambiguous ones ($\hat{\beta} = -0.06$, CrI: $[-0.13,0.00]$, $\Pr(\hat{\beta} > 0) = 0.03$). In the spillover region

Table 4.5: **Experiment 1: Results for total reading times (SOI stimuli)**. can = canonicity, el = elision.

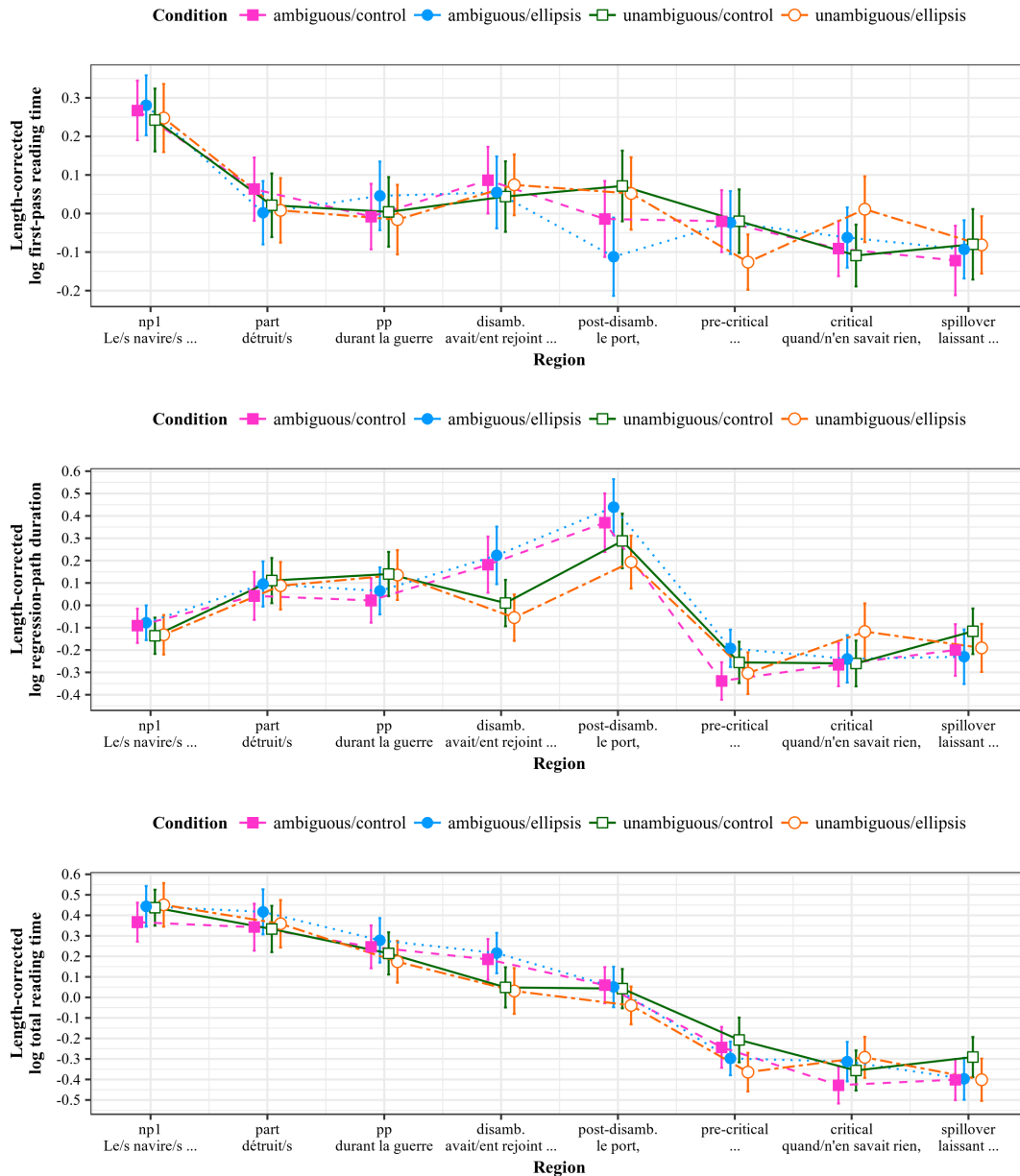
region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1_1 <i>Quel escrimeur</i>	el	0.01	-0.04	0.05	0.60
	can	0.04	0.00	0.08	0.96
	el × can	0.00	-0.04	0.04	0.51
np1_2 <i>de l'équipe nationale</i>	el	0.02	-0.02	0.06	0.83
	can	0.09	0.04	0.13	≈1.00
	el × can	0.03	-0.02	0.07	0.90
aux <i>a/ont salué</i>	el	0.02	-0.03	0.07	0.79
	can	0.14	0.08	0.20	≈1.00
	el × can	0.01	-0.04	0.05	0.61
np2_1 <i>les adversaires</i>	el	0.02	-0.03	0.06	0.80
	can	0.07	0.03	0.12	≈1.00
	el × can	-0.01	-0.05	0.04	0.40
np2_2 <i>du concours</i>	el	-0.03	-0.08	0.02	0.12
	can	0.04	0.00	0.09	0.98
	el × can	0.01	-0.03	0.05	0.69
crit <i>et quand,/avant la lutte,</i>	el	-0.04	-0.10	0.03	0.13
	can	0.00	-0.05	0.04	0.47
	el × can	0.00	-0.04	0.04	0.44
spillover <i>si je puis ...</i>	el	-0.04	-0.09	0.02	0.08
	can	0.00	-0.05	0.06	0.51
	el × can	0.00	-0.06	0.07	0.54

following the critical region, control sentences showed longer first-pass reading times than elided sentences ($\hat{\beta}=0.04$, CrI: $[-0.01,0.09]$, $\Pr(\hat{\beta} > 0)=0.95$).

Regression-path durations

Results for regression-path durations are shown in Table 4.8. At the disambiguating region, regression-path durations were longer in the ambiguous compared to the unambiguous conditions ($\hat{\beta} = 0.11$, CrI: $[0.03, 0.20]$, $\Pr(\hat{\beta} > 0) = 0.99$). The effect continued into the post-disambiguating region ($\hat{\beta}=0.08$, CrI: $[-0.01,0.17]$, $\Pr(\hat{\beta} > 0)=0.96$). At the pre-critical region, there was an interaction between elision and ambiguity ($\hat{\beta}=0.04$, CrI: $[-0.01,0.04]$, $\Pr(\hat{\beta} > 0) = 0.95$), such that ambiguity increased regression-path durations in the elision conditions while the opposite was observed in the control conditions, with neither effect driving the interaction.

Figure 4.2: **Reading measures by region for RRC stimuli (Experiment 1).** All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.



Total reading times

Total reading times at the disambiguating region were higher in the ambiguous than in the unambiguous conditions ($\hat{\beta} = 0.09$, CrI: [0.02, 0.15], $\Pr(\hat{\beta} > 0) \approx 1$). Table 4.9 shows results across all regions of interest.

Table 4.6: **Experiment 1: Results for question response accuracy and response times (RRC stimuli)**. amb = ambiguity, el = elision.

response accuracy (logit scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	0.29	-0.05	0.66	0.95
amb	0.08	-0.25	0.41	0.70
el \times amb	-0.18	-0.56	0.19	0.17
response time (log scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	-0.03	-0.10	0.04	0.22
amb	0.01	-0.05	0.07	0.59
el \times amb	0.03	-0.03	0.09	0.84

Results – Triple lexical ambiguity (TLA)

Length-corrected reading time measures by region of interest for the TLA stimuli are shown in Figure 4.3.

Comprehension probes

Results for question response accuracy and response latency are shown in Table 4.10. The mean probe response accuracy for TLA stimuli was 80%. There were no effects of the experimental manipulations on response accuracy or response times.

First-pass reading times

Table 4.11 shows the results for first-pass reading times. At the initial noun phrase, first-pass reading times were elevated in the ambiguous compared to the unambiguous conditions ($\hat{\beta} = 0.04$, CrI: [0.00, 0.08], $\Pr(\hat{\beta} > 0) = 0.96$). At the post-verbal region, there was an unexpected interaction ($\hat{\beta} = 0.04$, CrI: [0.00, 0.08], $\Pr(\hat{\beta} > 0) = 0.96$), which is likely to be spurious as elided and control sentences did not differ at this point. Speculatively, however, participants may have had preview of the critical region on the following line of text, even though this preview would have needed to cross two empty lines.

At the pre-critical region, first-pass reading times were shorter for ellipsis than for control sentences ($\hat{\beta} = -0.06$, CrI: [-0.12, -0.01], $\Pr(\hat{\beta} > 0) = 0.01$). In the spillover region following

Table 4.7: **Experiment 1: Results for first-pass reading times (RRC stimuli).** amb = ambiguity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s navire/s</i>	el	0.00	-0.05	0.05	0.57
	amb	0.01	-0.04	0.07	0.69
	el \times amb	0.00	-0.04	0.04	0.57
part <i>détruit/s</i>	el	-0.02	-0.06	0.03	0.21
	amb	0.01	-0.04	0.05	0.61
	el \times amb	-0.02	-0.07	0.04	0.27
pp <i>durant la guerre</i>	el	0.01	-0.05	0.07	0.63
	amb	0.01	-0.06	0.07	0.61
	el \times amb	0.02	-0.04	0.07	0.75
disamb <i>avait/ent rejoint</i>	el	0.00	-0.06	0.06	0.46
	amb	0.01	-0.05	0.06	0.58
	el \times amb	-0.02	-0.06	0.03	0.23
post-disamb <i>le port</i>	el	-0.02	-0.07	0.03	0.16
	amb	-0.06	-0.13	0.00	0.03
	el \times amb	-0.02	-0.08	0.04	0.27
pre-crit ...	el	-0.03	-0.10	0.04	0.23
	amb	0.02	-0.02	0.07	0.85
	el \times amb	0.02	-0.03	0.07	0.81
crit <i>quand,/n'en savait rien,</i>	el	0.00	-0.06	0.06	0.54
	amb	-0.02	-0.07	0.03	0.24
	el \times amb	0.00	-0.04	0.05	0.57
spillover <i>laissant</i>	el	0.04	-0.01	0.09	0.95
	amb	-0.02	-0.06	0.03	0.22
	el \times amb	-0.02	-0.07	0.02	0.15

the critical region, there was an effect of ambiguity on first-pass reading times, such that reading times were elevated in the ambiguous conditions ($\hat{\beta} = 0.04$, CrI: [0.00,0.08], $\Pr(\hat{\beta} > 0) = 0.97$).

Regression-path durations

Regression-path durations were elevated for ambiguous compared to unambiguous sentences on the first noun phrase ($\hat{\beta} = 0.04$, CrI: [0.01,0.10], $\Pr(\hat{\beta} > 0) = 0.95$). Longer regression-path durations were also observed for ambiguous sentences at the adjective ($\hat{\beta} = 0.11$, CrI: [0.04,0.18], $\Pr(\hat{\beta} > 0) \approx 1$). Next, there was a surprising interaction between ambiguity and

Table 4.8: **Experiment 1: Results for regression-path durations (RRC stimuli)**. amb = ambiguity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s navire/s</i>		See first-pass reading times			
part <i>détruit/s</i>	el	0.00	-0.06	0.06	0.56
	amb	-0.02	-0.11	0.07	0.31
	el × amb	0.01	-0.04	0.07	0.67
pp <i>durant la guerre</i>	el	0.01	-0.05	0.07	0.63
	amb	-0.05	-0.13	0.03	0.11
	el × amb	0.01	-0.05	0.07	0.65
disamb <i>avait/ent rejoint</i>	el	-0.01	-0.08	0.06	0.40
	amb	0.11	0.03	0.20	0.99
	el × amb	0.02	-0.05	0.09	0.69
post-disamb <i>le port</i>	el	-0.01	-0.08	0.07	0.44
	amb	0.08	-0.01	0.17	0.96
	el × amb	0.04	-0.02	0.10	0.88
pre-crit ...	el	0.02	-0.05	0.09	0.73
	amb	0.01	-0.05	0.07	0.61
	el × amb	0.04	-0.01	0.09	0.95
crit <i>quand,/n'en savait rien,</i>	el	-0.04	-0.11	0.03	0.11
	amb	-0.03	-0.10	0.03	0.17
	el × amb	0.01	-0.05	0.07	0.58
spillover <i>laissant</i>	el	0.04	-0.03	0.11	0.88
	amb	-0.03	-0.09	0.03	0.13
	el × amb	-0.03	-0.09	0.03	0.17

elision in regression-path durations on the verb+pronoun region ($\hat{\beta} = 0.07$, CrI: [0.01,0.12], $\Pr(\hat{\beta} > 0) = 0.99$). Elision and control sentences did not differ at this point, and planned comparisons revealed the interaction to be driven by neither sentence type, leading us to assume that the effect is spurious. At the pre-critical region, ambiguous sentences had shorter regression paths than unambiguous sentences ($\hat{\beta} = -0.04$, CrI: [-0.09,0.01], $\Pr(\hat{\beta} > 0) = 0.05$). At the critical region, regression-path durations were shorter for ellipsis than for control sentences ($\hat{\beta} = -0.06$, CrI: [-0.13,0.00], $\Pr(\hat{\beta} > 0) = 0.02$). Table 4.12 lists the results for all regions of interest.

Table 4.9: **Experiment 1: Results for total reading times (RRC stimuli)**. amb = ambiguity, el = elision.

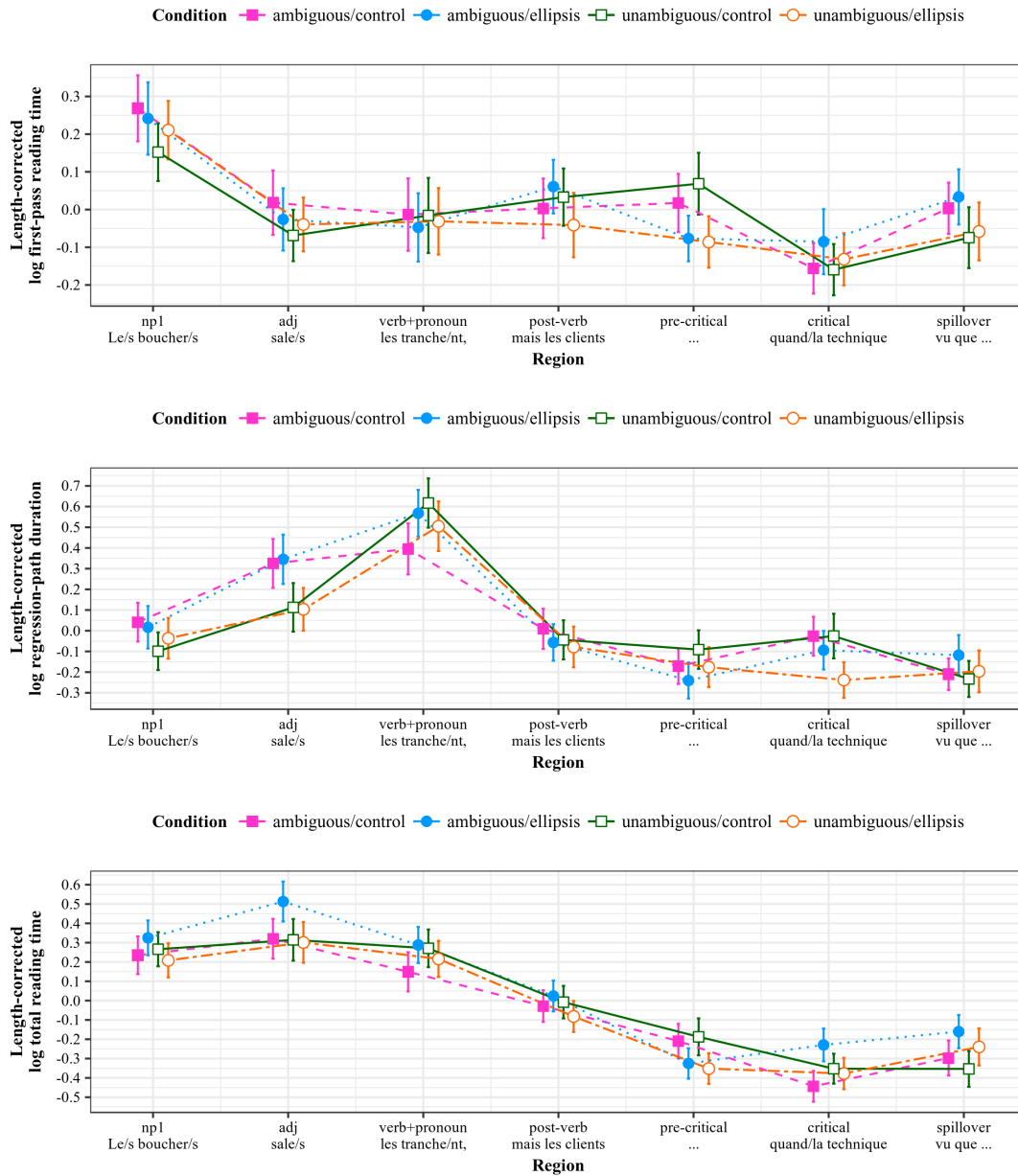
region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s navire/s</i>	el	0.02	-0.03	0.07	0.81
	amb	-0.01	-0.06	0.04	0.32
	el × amb	0.01	-0.03	0.06	0.72
part <i>détruit/s</i>	el	0.02	-0.03	0.08	0.83
	amb	0.03	-0.03	0.08	0.83
	el × amb	0.01	-0.04	0.06	0.66
pp <i>durant la guerre</i>	el	0.00	-0.05	0.05	0.51
	amb	0.03	-0.01	0.08	0.94
	el × amb	0.01	-0.03	0.06	0.73
disamb <i>avait/ent rejoint</i>	el	0.00	-0.05	0.05	0.55
	amb	0.09	0.02	0.15	≈1.00
	el × amb	0.01	-0.06	0.08	0.62
post-diamb <i>le port</i>	el	-0.02	-0.07	0.03	0.21
	amb	0.02	-0.04	0.09	0.80
	el × amb	0.02	-0.03	0.07	0.76
pre-crit ...	el	-0.05	-0.13	0.02	0.08
	amb	0.00	-0.04	0.05	0.57
	el × amb	0.02	-0.02	0.07	0.84
crit <i>quand,/n'en savait rien,</i>	el	0.01	-0.09	0.12	0.61
	amb	-0.03	-0.09	0.02	0.14
	el × amb	0.03	-0.02	0.08	0.85
spillover <i>laissant</i>	el	0.04	-0.03	0.10	0.89
	amb	-0.03	-0.07	0.02	0.12
	el × amb	0.00	-0.04	0.05	0.57

Total reading times

Results for total reading times are listed in Table 4.13. The adjective showed an interaction between elision and ambiguity ($\hat{\beta}=0.05$, CrI: [0.00,0.09], $\Pr(\hat{\beta} > 0)=0.98$), driven by longer reading times in ambiguous elided sentences ($\hat{\beta}=0.07$, CrI: [-0.02,0.16], $\Pr(\hat{\beta} > 0)=0.95$). An interaction with the same sign was evident at the verb+pronoun region ($\hat{\beta}=0.05$, CrI: [0.01,0.09], $\Pr(\hat{\beta} > 0)=0.99$), but was driven by neither pair of conditions.

At the pre-critical region, sentences with elision showed shorter total reading times than control sentences ($\hat{\beta} = -0.06$, CrI: [-0.13,0.00], $\Pr(\hat{\beta} > 0) = 0.03$). At the critical region, elided sentences showed longer reading times than control sentences ($\hat{\beta} = 0.05$, CrI:

Figure 4.3: Reading measures by region for TLA stimuli (Experiment 1). All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.



$[-0.01, 0.11]$, $\Pr(\hat{\beta} > 0) = 0.96$). There was also an interaction between ellipsis and ambiguity ($\hat{\beta} = 0.06$, CrI: $[0.01, 0.10]$, $\Pr(\hat{\beta} > 0) = 0.99$), mainly driven by longer reading times in ambiguous elided sentences ($\hat{\beta} = 0.07$, CrI: $[0.01, 0.13]$, $\Pr(\hat{\beta} > 0) = 0.98$). At the spillover

Table 4.10: **Experiment 1: Results for question response accuracy and response times (TLA stimuli).** amb = ambiguity, el = elision.

response accuracy (logit scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	-0.13	-0.56	0.29	0.28
amb	-0.07	-0.34	0.19	0.30
el \times amb	0.00	-0.39	0.38	0.52
response time (log scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
el	0.00	-0.04	0.04	0.53
amb	0.00	-0.04	0.05	0.57
el \times amb	0.01	-0.04	0.05	0.64

region following the critical region, elided sentences again showed longer total reading times than non-elided sentences ($\hat{\beta}=0.06$, CrI: $[-0.01,0.13]$, $\Pr(\hat{\beta} > 0)=0.96$).

Discussion

Garden paths in French

Our design rested on the assumption that all three stimulus types would cause garden-path effects, which was largely borne out in the data. Off-line acceptability judgments had already shown that the ambiguous and non-canonical versions of TLA and SOI stimuli, respectively, received lower ratings, which we took to indicate processing difficulty on part of our participants. RRC stimuli, for which no effect had been found in off-line acceptability judgments, nevertheless showed an effect of ambiguity in on-line measures, specifically in both regression-path durations and total reading times at the disambiguating finite verb.

For SOI stimuli, on-line processing difficulty was also observed at the disambiguating region, again in the form of longer regression paths and higher total reading times. The garden-path effect for the OVS sentences appears to be quite strong, as evidenced by its persistence across the following regions and the effect on total reading times across the whole antecedent clause. It is also striking that SOI stimuli showed the lowest comprehension accuracy among the three item types. This, however, may be due to the fact that as SOI sentences contained less material between antecedent and ellipsis site than both RRC and TLA sentences, a

Table 4.11: **Experiment 1: Results for first-pass reading times (TLA stimuli).** amb = ambiguity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s boucher/s</i>	el	0.01	-0.04	0.05	0.62
	amb	0.04	0.00	0.08	0.96
	el × amb	-0.02	-0.06	0.02	0.17
adj <i>sale/s</i>	el	0.00	-0.05	0.04	0.47
	amb	0.02	-0.02	0.06	0.82
	el × amb	-0.02	-0.07	0.02	0.14
verb+pronoun <i>les tranche/nt,</i>	el	-0.01	-0.06	0.03	0.26
	amb	-0.01	-0.06	0.05	0.38
	el × amb	0.00	-0.08	0.06	0.45
post-verb <i>mais les clients</i>	el	0.00	-0.04	0.04	0.45
	amb	0.02	-0.02	0.05	0.80
	el × amb	0.04	0.00	0.08	0.96
pre-crit ...	el	-0.06	-0.12	-0.01	0.01
	amb	-0.01	-0.05	0.03	0.24
	el × amb	0.01	-0.03	0.05	0.70
crit <i>quand,/la technique,</i>	el	0.03	-0.03	0.09	0.88
	amb	0.01	-0.03	0.05	0.62
	el × amb	0.01	-0.03	0.05	0.72
spillover <i>vu que ...</i>	el	0.01	-0.02	0.05	0.79
	amb	0.04	0.00	0.08	0.97
	el × amb	0.00	-0.04	0.04	0.45

larger amount of comprehension questions necessarily targeted the assignment of argument roles within the antecedent, so that misinterpretations could be more easily detected.

The reactivation hypothesis

The current study yielded no evidence in favor of the reactivation hypothesis, which predicted that an ellipsis should be easier to process in sentences where the antecedent has been syntactically reanalyzed. For SOI stimuli, we found no evidence that having experienced a garden path while processing the antecedent had any effect on the subsequent processing of the ellipsis. For TLA stimuli, we observed the opposite of what the reactivation hypothesis predicted: when the antecedent clause was temporarily ambiguous, ellipsis was more difficult to process. For RRC stimuli, there was evidence of increased processing difficulty in the pre-critical region when the antecedent was temporarily ambiguous and the

Table 4.12: **Experiment 1: Results for regression-path durations (TLA stimuli)**. amb = ambiguity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s boucher/s</i>	el	0.01	-0.04	0.07	0.68
	amb	0.04	-0.01	0.10	0.95
	el × amb	-0.02	-0.07	0.03	0.20
adj <i>sale/s</i>	el	0.00	-0.05	0.06	0.56
	amb	0.11	0.04	0.18	≈1.00
	el × amb	0.00	-0.07	0.07	0.54
verb+pronoun <i>les tranche/nt,</i>	el	0.02	-0.05	0.08	0.69
	amb	-0.05	-0.15	0.06	0.20
	el × amb	0.07	0.01	0.12	0.99
post-verb <i>mais les clients</i>	el	-0.02	-0.07	0.03	0.18
	amb	0.02	-0.04	0.07	0.71
	el × amb	0.00	-0.06	0.05	0.43
pre-crit ...	el	-0.05	-0.11	0.02	0.09
	amb	-0.04	-0.09	0.01	0.05
	el × amb	0.00	-0.05	0.05	0.53
crit <i>quand,/la technique,</i>	el	-0.06	-0.13	0.00	0.02
	amb	0.03	-0.03	0.08	0.84
	el × amb	0.03	-0.02	0.09	0.91
spillover <i>vu que ...</i>	el	0.03	-0.01	0.09	0.92
	amb	0.02	-0.03	0.07	0.80
	el × amb	0.00	-0.05	0.06	0.58

following critical region contained an ellipsis; this effect, however, is somewhat difficult to gauge as it could stem either from parafoveal preview of the ellipsis site or from reading a pronoun in the pre-critical region in the control conditions, or both.

The first interesting question to ask is why German readers showed a reactivation effect in Paape (2016) while the French readers showed the opposite effect, namely difficulty processing the ellipsis in the presence of temporarily ambiguous antecedents in TLA stimuli, and possibly RRC stimuli as well. A second question is why French readers showed a clear reverse effect only with TLA stimuli, but not in sentences with subject-object inversion or when the antecedent contained a reduced relative clause, despite evidence that there was garden-pathing in these latter stimuli. We believe that the difference in methods between Paape’s (2016) experiment and the current study, the idiosyncrasies of the stimuli used

Table 4.13: **Experiment 1: Results for total reading times (TLA stimuli).** amb = ambiguity, el = elision.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>Le/s boucher/s</i>	el	0.01	-0.03	0.06	0.69
	amb	0.03	-0.02	0.08	0.89
	el × amb	0.03	-0.01	0.08	0.93
adj <i>sale/s</i>	el	0.05	-0.02	0.12	0.92
	amb	0.06	-0.02	0.14	0.94
	el × amb	0.05	0.00	0.09	0.98
verb+pronoun <i>les tranche/nt,</i>	el	0.02	-0.03	0.08	0.79
	amb	-0.01	-0.07	0.06	0.41
	el × amb	0.05	0.01	0.09	0.99
post-verb <i>mais les clients</i>	el	0.00	-0.06	0.05	0.44
	amb	0.02	-0.03	0.06	0.77
	el × amb	0.03	-0.02	0.08	0.91
pre-crit ...	el	-0.06	-0.13	0.00	0.03
	amb	-0.01	-0.05	0.04	0.38
	el × amb	0.01	-0.03	0.06	0.71
crit <i>quand,/la technique,</i>	el	0.05	-0.01	0.11	0.96
	amb	0.01	-0.04	0.05	0.64
	el × amb	0.06	0.01	0.10	0.99
spillover <i>vu que ...</i>	el	0.06	-0.01	0.13	0.96
	amb	0.03	-0.02	0.07	0.89
	el × amb	0.00	-0.04	0.05	0.59

as well as cross-linguistic differences in the acceptability of object-initial sentences may be jointly responsible for the diverging results. We discuss these points next.

While we chose to use eye tracking to make processing of the stimuli more natural for our participants, different styles and strategies may emerge in comparison to the self-paced reading methodology used in the original study. In non-cumulative self-paced reading, participants are aware that they cannot make regressions to earlier parts of the sentence, which could conceivably result in a more careful processing strategy. By ‘more careful’, we mean a strategy that minimizes strategic underspecification on part of the reader: a reader who knows that regressions are possible might be tempted to not fully disambiguate the syntactic structure of the antecedent, hoping that later information would point towards the correct analysis, or that choosing the correct analysis was simply not relevant to the task at hand. While evidence for underspecification has repeatedly been found in self-paced reading

experiments (Merchant, 2013; Swets et al., 2008), its use may be still more prevalent when readers know that they can re-read earlier parts of the sentence at their leisure.

Especially with regard to our TLA stimuli, we suspect that at least on some trials, participants failed or neglected to resolve the ambiguity before proceeding past the antecedent clause. The first piece of evidence for this assumption is that TLA stimuli failed to show a clear garden-path effect, that is, an indication of processing difficulty for ambiguous sentences at or directly after the point of disambiguation. Instead, an ambiguity disadvantage was visible at the adjective, likely indicating that participants were torn between the two possible readings, that is, a competition effect. If participants failed to settle on a reading, it is not surprising that there was no evidence of revision at the point of disambiguation. Looking at first-pass reading times and regression-path durations, there was even a trend for ambiguous sentences to be easier to process than unambiguous ones at this position. We assume that when participants reached the ellipsis site, they realized that they did not have a suitable antecedent available for retrieval, which led to elevated total reading times at this position (as well as regressions towards the antecedent, as shown in the next section). No reactivation advantage was observed because reanalysis sometimes did not take place prior to reaching the ellipsis site.

With regard to our SOI stimuli, we have some anecdotal evidence from participants who claimed that the non-canonical versions were ungrammatical. While we have it on good authority that OVS word order in *wh*-questions is, in fact, grammatical in French (Lahousse, 2003), it is nevertheless worrying that some subjects would claim otherwise, which may indicate that they failed to parse the non-canonical stimuli. If a large enough group of participants was under the impression of being presented with an ungrammatical sentence in these cases, it would not be surprising if interpretation of the ellipsis was largely suspended in non-canonical sentences, given that the clause could not be assigned a meaning anyway.

It is possible that the impression of ungrammaticality was partly due to an experimenter effect: as the main author, who carried out the experiment, is not a native speaker of French, participants may have been biased to expect some proportion of ungrammatical sentences. As a precaution, we ensured that a second experimenter who was a native speaker was always present during the experimental sessions, and explicitly informed participants that the stimuli had been designed by a native speaker. Still, some participants insisted on having read ungrammatical sentences even when the second experimenter explained the construction during debriefing.

There were no comments indicating that RRC stimuli were perceived as ungrammatical. Rather, it appears that they were the easiest to process, compared to the other stimulus types. We thus speculate that the reactivation effect for RRC stimuli was too subtle to be clearly

detectable in our experiment, possibly because the construction causes garden-pathing less reliably in French, as suggested by the results of our pre-study. In this context, recall that our pre-test failed to show an effect of the ambiguity manipulation on acceptability ratings.

Exploratory analyses

Exploratory analysis I – Regressions and antecedent rereading in TLA stimuli

For TLA stimuli, the analyses of total reading times for the verb+pronoun region showed effects of the elision manipulation, possibly indicating differential amounts of rereading. However, based on this result alone, we cannot be sure when the rereading occurred: a pronoun appearing before the critical region may trigger regressions, leading to higher rereading times for the antecedent on a given trial even though the ellipsis has not been processed yet. To address this issue, we separated the data for TLA stimuli into fixations made before the critical region was first fixated on a given trial and fixations made after the first fixation on the critical region and computed reading measures separately using the `em2` R package (Logačev & Vasishth, 2013).

When only reading times that were generated before the first fixation on the critical region were considered, none of the three regions of the antecedent clause, that is, the initial noun phrase, the adjective and the verb+pronoun region, showed any effects of the experimental manipulations. The same was true when only reading times that were generated *after* the first fixation on the critical region were considered.

Looking at overall rereading probabilities for the antecedent regions, there was an interaction at the first noun phrase ($\hat{\beta} = 0.25$, CrI: [0.01, 0.49], $\Pr(\hat{\beta} > 0) = 0.98$) as well as at the adjective ($\hat{\beta} = 0.22$, CrI: [0.00, 0.45], $\Pr(\hat{\beta} > 0) = 0.97$), the latter being driven by additional refixations in ambiguous sentences with elision ($\hat{\beta} = 0.32$, CrI: [0.03, 0.63], $\Pr(\hat{\beta} > 0) = 0.99$).

When the rereading times for all antecedent regions after the first fixation on the critical region were summed up, there was an effect of ambiguity ($\hat{\beta} = 0.13$, CrI: [-0.01, 0.26], $\Pr(\hat{\beta} > 0) = 0.97$), such that ambiguous sentences caused more rereading.

Discussion

The results of the exploratory study indicate that the adjective region was reread more often in elided sentences with ambiguous antecedents, a finding that matches the increased reading times at the critical region in this condition: when readers realized that they could not retrieve a suitable antecedent structure, they presumably returned to the antecedent in order to attempt a reparse. Furthermore, the increase in antecedent rereading times in the ambiguous conditions may indicate that both ellipsis and pronouns were affected by the occasional unavailability of an antecedent, which led to attempts to reparse the problematic earlier part of the sentence.

Exploratory analysis II – Effects of lexical bias in TLA stimuli

There was no evidence for a classical garden path in TLA stimuli, as no effect of ambiguity was detected in the disambiguating region, even though ambiguous sentences generated longer regression-path durations on the adjective. In order to make sense of this finding, it is important to consider that the word *sale*, ‘dirty’, which shares its form with the third-person singular of *saler*, ‘to salt’, will be subject to lexical competition: on the one hand, the syntactic context suggests a simple SVO sentence, which favors the finite verb reading. However, the French Wikipedia corpus available at <http://corp.hum.sdu.dk/cqp.fr.html> (37.8 m words; Bick, 2005) shows the form *sale* to appear as an adjective 160 times, compared to only 9 occurrences as a finite verb. This implies that at least for the form *sale*, lexical frequency favors the adjective reading while the syntactic context favored the verb reading.

In order to further explore the influence of lexical competition on the ambiguity effect in TLA stimuli, we extracted token counts for both the verb and adjective readings of the relevant forms appearing in our items from the aforementioned Wikipedia corpus and computed the log ratio as a measure of bias toward the verb reading.

When entered as a covariate into the statistical analysis, the bias predictor showed no effect on first-pass reading times or regression-path durations for the adjective. In total reading times, in addition to a main effect of elision ($\hat{\beta} = 0.06$, CrI: [0.00, 0.11], $\Pr(\hat{\beta} > 0) = 0.97$) and a two-way interaction between ambiguity and elision ($\hat{\beta} = 0.05$, CrI: [-0.01, 0.10], $\Pr(\hat{\beta} > 0) = 0.96$), there was a two-way interaction between ambiguity and bias ($\hat{\beta} = -0.03$, CrI: [-0.06, 0.00], $\Pr(\hat{\beta} > 0) = 0.02$), such that a bias towards the verb reading led to longer reading times in the unambiguous conditions with adjective morphology only ($\hat{\beta} = 0.06$, CrI: [0.00, 0.12], $\Pr(\hat{\beta} > 0) = 0.97$).

When only total reading times prior to fixating the critical region were taken into account, there was evidence of a two-way interaction between bias and ambiguity ($\hat{\beta} = -0.03$, CrI: $[-0.06, -0.01]$, $\Pr(\hat{\beta} > 0) = 0.01$), again driven by an effect of bias in the unambiguous conditions, such that verb-biased participles showed longer reading times in the presence of adjective morphology ($\hat{\beta} = 0.05$, CrI: $[0.00, 0.11]$, $\Pr(\hat{\beta} > 0) = 0.98$). There were no effects of bias when only reading times after the first fixation on the critical region were taken into account.

Discussion

While ambiguous adjectives were more difficult to process overall, most likely due to increased lexical competition, it appears that participants perceived a clash between the lexical bias of the root form and the morphological adjective marker in the unambiguous conditions, which also caused processing difficulty. As there was no interaction with the elision manipulation and no evidence that lexical bias interacted with the experimental factors after participants had read the critical region, it appears that lexical bias probably exerted an effect on the initial analysis of the antecedent, but not on later attempts at reparsing the structure.

Experiment 2

Experiment 1 yielded no evidence in favor of the reactivation hypothesis, which casts doubt on whether the results of Paape (2016) are reproducible across different experimental paradigms and languages. In order to get a clearer picture, our second eye-tracking study attempted to replicate the original result with the same materials and roughly the same number of participants, but using eye tracking instead of self-paced reading. If the eye-tracking version of the German study yields comparable results, this would strengthen the claim that the pattern observed in the original experiment was not a mere artifact, and suggest that the diverging results in French are due to differences in the constructions investigated and not between the methods used.

Materials

The experimental materials for Experiment 2 were taken directly from Paape (2016) (see Appendix I). Regions of interest were the same as in the original study. The factors

word order (SVO vs. OVS) and case marking on the initial noun phrase (ambiguous vs. unambiguous) were manipulated in a 2×2 design, as shown below. As explained in the introduction, feminine nouns in German are not overtly marked for case, meaning that the first noun phrase of a main clause could either be the subject or the object, with the subject reading being preferred (Hemforth, 1993; Meng & Bader, 2000). The finite auxiliary *hatte/n*, ‘had’, agrees in number with either the first noun phrase, indicating that the word order is SVO, or with the second noun phrase, indicating that the word order is OVS.

(29) a. **Ambiguous / OVS**

Eine Sympathisantin
 A sympathizer.fem.**nom/acc**
 der Opposition_{np1} ◇ hatten_{aux} ◇ die Rebellen_{np2} ◇ ...
 of the opposition **had.pl** the rebels.nom/acc

b. **Ambiguous / SVO**

Eine Sympathisantin
 A sympathizer.fem.**nom/acc**
 der Opposition_{np1} ◇ hatte_{aux} ◇ die Rebellen_{np2} ◇ ...
 of the opposition **had.sg** the rebels.nom/acc

c. **Unambiguous / OVS**

Einen Sympathisanten
 A sympathizer.masc.**acc**
 der Opposition_{np1} ◇ hatten_{aux} ◇ die Rebellen_{np2} ◇ ...
 of the opposition **had.pl** the rebels.nom/acc

d. **Unambiguous / SVO**

Ein Sympathisant
 A sympathizer.masc.**nom**
 der Opposition_{np1} ◇ hatte_{aux} ◇ die Rebellen_{np2} ◇ ...
 of the opposition **had.sg** the rebels.nom/acc
 ... laut einem Bericht_{adj} ◇ maßgeblich unterstützt_{vp}
 according to a report decisively supported
 ◇ aber ◇ die Regierung ◇ konnte ◇ nicht ◇ nachweisen, ◇ wie_{critical}
 but the government could not substantiate **how**
 ◇ so sehr ◇ sich ◇ die Untersuchungskommission ◇ auch ◇ bemühte.
 so greatly SELF the investigative commission too struggled

‘The rebels had supported a sympathizer (OVS, a/c) / A sympathizer had supported the rebels (SVO, b/d), but the government could not substantiate how, no matter how hard the investigative commission tried.’

Sentences (29c,d) are control conditions in which the initial noun phrase is overtly marked for case, so that no additional processing difficulty is expected when the auxiliary agrees

with the second noun phrase in (29c). As in the original study, but unlike in our Experiment 1, there were no control conditions without ellipsis in Experiment 2.

Predictions

If the change from self-paced reading to eye tracking does not have an influence on participants' reading strategies, we expect to find the same pattern of results as in the original study of Paape (2016). First, a garden-path effect is expected at the auxiliary or at the second noun phrase, which should appear in the form of an interaction: in the original study, OVS sentences were especially difficult to process when the initial noun phrase carried ambiguous case marking, suggesting that readers misanalyzed this noun phrase as being the subject at first.

At or directly after the critical ellipsis region, OVS sentences should be easier to process than SVO sentences in the ambiguous as compared to the unambiguous conditions, given that the reanalysis of the antecedent structure should lead to reactivation of the memory trace. In principle, this pattern may be visible in any of the eye-tracking measures analyzed (first-pass reading time, regression-path duration, total reading time), as reading times from self-paced reading do not allow inferences as to whether an effect arises early or late in the processing stream.

Furthermore, Paape (2016) observed the same interaction directly before the ellipsis site, which was interpreted as evidence for predictive processing: participants were argued to maintain an expectation of the upcoming ellipsis gap, effectively trying to pre-fill the gap before having encountered it. Paape also observed a processing disadvantage for OVS sentences in one of the ellipsis spillover regions, as well as a numerical trend towards the same disadvantage both directly before and after the ellipsis site. This was interpreted as an effect of the mismatch between the OVS antecedent and the ellipsis gap, which sets a cue for an SOV antecedent (see introduction).

It is possible that the relative freedom of the eye-tracking paradigm as compared to self-paced reading, especially with regard to the possibility of regressions, may change the pattern of results. For instance, readers may choose not to resolve the ambiguity in the initial clause, like the French participants did for TLA stimuli in Experiment 1, and experience processing difficulty at the ellipsis site because they do not have a fully analyzed representation of the antecedent available for retrieval. Such a result would not constitute evidence against the reactivation hypothesis *per se*, but would highlight the impact of different reading strategies on the probability of a correct retrieval at the critical region.

Participants

We recruited 62 German native speakers as subjects. All had normal or corrected-to-normal eyesight and were either paid 7 € or received course credit for their participation. Informed consent was obtained from all participants prior to the experiment.

Procedure

The procedure was analogous to that of Experiment 1, with the following changes. Eye movements were recorded by an SR Research Eyelink 1000 tracker with a desktop-mounted camera setup. Sentences were presented in 21 pt Arial font at a resolution of 1680×1050 pixels, with participants sitting at a distance of 61 cm from the presentation screen. The screen was 83.8 cm wide, so that each character displayed equaled 0.98° of visual angle. Recalibration of the eye-tracker was performed as needed when tracking accuracy fell below acceptable levels. Participants signaled completion of a trial by looking in the lower right corner of the screen for one second. Unlike in Experiment 1, filler sentences as well as experimental sentences were always followed by one of two types of comprehension test. The first type consisted of a statement that had to be judged as being either true or false while the second type required participants to fill in a gap within a statement, choosing one out of four options, as in Paape (2016). As in the original study, a subset of comprehension probes targeted the critical wh-pronoun while another subset targeted the argument structure of the antecedent.

Data Analysis

Data analysis was carried out analogously to Experiment 1. Again, the experimental factors were sum-coded, with ambiguous and OVS conditions coded as 1 and unambiguous and SVO conditions coded as -1 , respectively. As in Experiment 1, a log transformation was applied to the dependent measures prior to analysis, and the log-transformed length of each region in letters was entered into the analysis as a covariate. Due to experimenter error, question response data from 10 participants was lost, thus the reported accuracy and response latency results refer to the remaining 52 participants.

Results

Length-corrected reading time measures by region of interest for Experiment 2 are shown in Figure 4.4.

Comprehension probes

Results for probe response accuracy and response latency are shown in Table 4.14. The overall probe response accuracy was 85%. Participants reached 92% accuracy when asked to supply the critical wh-pronoun and 76% accuracy when the argument structure of the antecedent was probed. There were no effects of the experimental manipulations on response times. For response accuracy, there was an effect of word order, such that probes about OVS sentences were answered less accurately than probes about SVO sentences ($\hat{\beta} = -0.25$, CrI: $[-0.53, 0.01]$, $\Pr(\hat{\beta} > 0) = 0.03$).

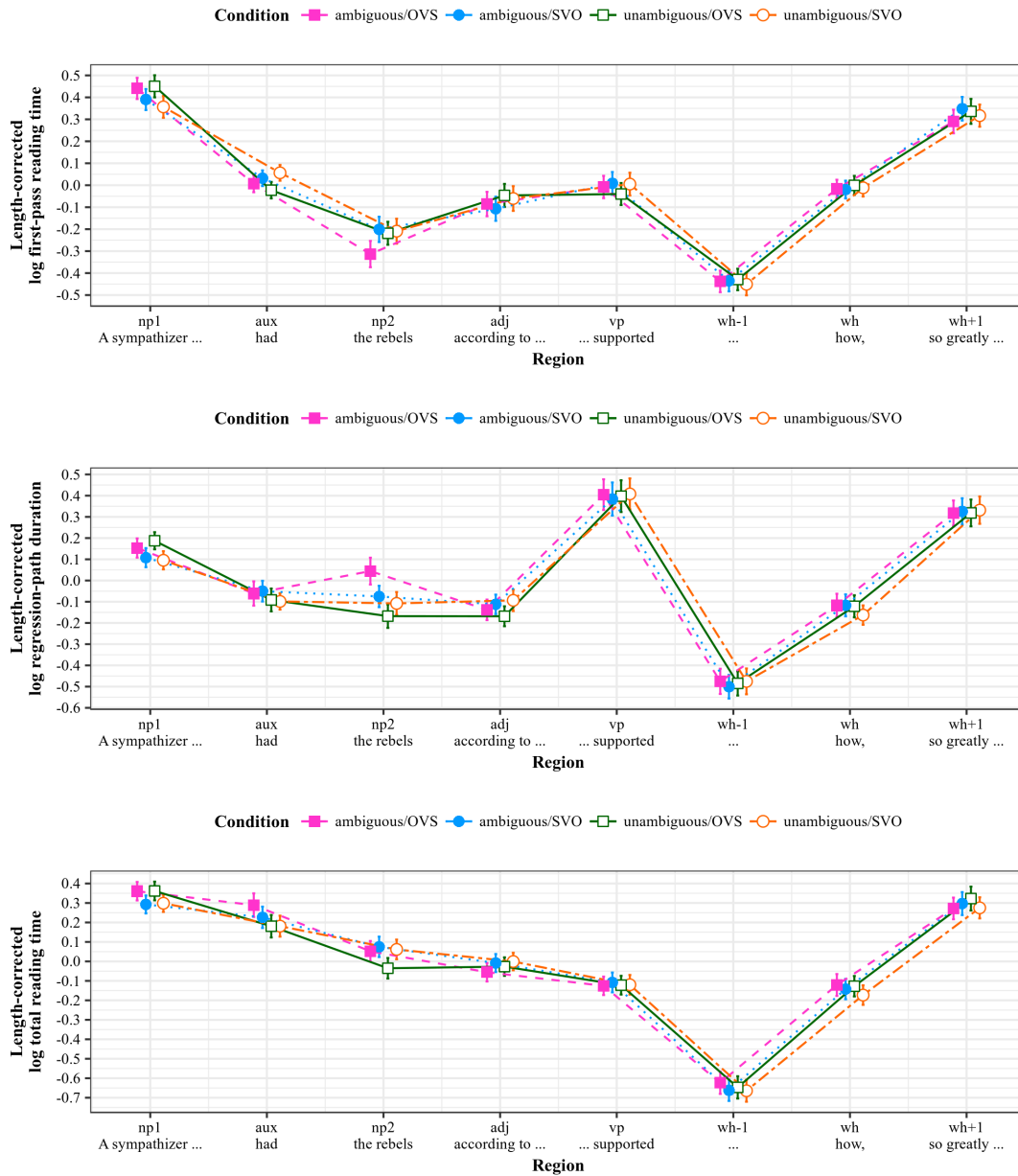
Table 4.14: **Experiment 2: Results for question response accuracy and response times.** case = case marking, ord = word order.

response accuracy (logit scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
case	-0.05	-0.30	0.21	0.35
ord	-0.25	-0.53	0.01	0.03
case \times ord	-0.19	-0.44	0.07	0.07
response time (log scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
case	0.00	-0.03	0.02	0.46
ord	0.01	-0.02	0.03	0.64
case \times ord	0.01	-0.01	0.03	0.77

In order to check for possible parsing failures, which should result in comprehension probes being answered incorrectly, we ran an additional analysis in which the type of the comprehension probe was added to the accuracy and response time models. Argument structure probes were coded as 1 and other probes as -1. The results of this ancillary analysis are listed in Table 4.15.

Response time only showed an effect of probe type, such that argument structure probes were responded to more quickly than the other probe types ($\hat{\beta} = -0.11$, CrI: $[-0.19, -0.04]$, $\Pr(\hat{\beta} > 0) \approx 0$). With regard to accuracy, in addition to main effects of word order

Figure 4.4: **Reading measures by region (Experiment 2)**. All measures log-transformed and residualized against region length in characters; error bars show 95% intervals.



($\hat{\beta} = -0.26$, CrI: $[-0.51, -0.01]$, $\Pr(\hat{\beta} > 0) = 0.02$), probe type ($\hat{\beta} = -0.86$, CrI: $[-1.24, -0.50]$, $\Pr(\hat{\beta} > 0) \approx 0$), along with all two-way interactions (case marking \times word order: $\hat{\beta} = -0.24$, CrI: $[-0.47, -0.01]$, $\Pr(\hat{\beta} > 0) = 0.02$, word order \times probe type: $\hat{\beta} = -0.36$, CrI: $[-0.60, -0.12]$,

$\Pr(\hat{\beta} > 0) \approx 0$, case marking \times probe type: $\hat{\beta} = -0.29$, CrI: $[-0.53, -0.06]$, $\Pr(\hat{\beta} > 0) = 0.01$), there was also a three-way interaction ($\hat{\beta} = -0.26$, CrI: $[-0.50, -0.03]$, $\Pr(\hat{\beta} > 0) = 0.02$), such that ambiguous OVS sentences caused a much steeper drop in accuracy with argument structure probes than with the other probe types (56% vs. 92% mean accuracy).

Table 4.15: **Experiment 2: Results for question response accuracy by probe type.** case = case marking, ord = word order, pt = probe type.

response accuracy (logit scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
case	-0.08	-0.33	0.16	0.24
ord	-0.26	-0.51	-0.01	0.02
pt	-0.86	-1.24	-0.50	≈ 0.00
case \times ord	-0.24	-0.47	-0.01	0.02
case \times pt	-0.29	-0.53	-0.06	0.01
ord \times pt	-0.36	-0.60	-0.12	≈ 0.00
case \times ord \times pt	-0.26	-0.50	-0.03	0.02
response time (log scale)				
parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
case	0.00	-0.03	0.02	0.45
ord	0.00	-0.03	0.04	0.59
pt	-0.11	-0.19	-0.04	≈ 0.00
case \times ord	0.01	-0.02	0.03	0.73
case \times pt	0.00	-0.02	0.03	0.56
ord \times pt	-0.01	-0.04	0.02	0.34
case \times ord \times pt	0.00	-0.02	0.02	0.38

First-pass reading times

First-pass reading times were longer for OVS compared to SVO sentences on the initial noun phrase ($\hat{\beta} = 0.04$, CrI: $[0.01, 0.06]$, $\Pr(\hat{\beta} > 0) \approx 1$), an effect that was reversed on the auxiliary ($\hat{\beta} = -0.04$, CrI: $[-0.06, -0.01]$, $\Pr(\hat{\beta} > 0) \approx 0$). The effect on the initial noun phrase is surprising given that there was no interaction with case marking, but may in part be due to parafoveal processing and/or misattributed fixations on the adjacent auxiliary, given that region's relatively small size. The adjunct phrase showed shorter first-pass reading times in the ambiguous compared to the unambiguous conditions ($\hat{\beta} = -0.02$, CrI: $[-0.05, 0.00]$, $\Pr(\hat{\beta} > 0) = 0.03$).

An interaction between word order and case marking became evident in the ellipsis spillover region ($\hat{\beta} = -0.02$, CrI: $[-0.05, 0.00]$, $\Pr(\hat{\beta} > 0) = 0.34$), driven by shorter reading times in the ambiguous OVS compared to the ambiguous SVO condition ($\hat{\beta} = -0.04$, CrI: $[-0.08, 0.00]$, $\Pr(\hat{\beta} > 0) = 0.02$). Table 4.16 shows parameter estimates across all regions of interest.

Table 4.16: **Experiment 2: Results for first-pass reading times.** case = case marking, ord = word order.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1	case	0.01	-0.02	0.03	0.67
<i>A sympathizer ...</i>	ord	0.04	0.01	0.06	≈ 1.00
	case \times ord	-0.01	-0.04	0.01	0.19
	case	0.01	-0.02	0.03	0.72
aux <i>had</i>	ord	-0.04	-0.06	-0.01	≈ 0.00
	case \times ord	0.02	-0.01	0.04	0.93
	case	-0.02	-0.05	0.01	0.06
np2 <i>the rebels</i>	ord	-0.03	-0.07	0.01	0.08
	case \times ord	-0.02	-0.05	0.01	0.06
	case	-0.02	-0.05	0.00	0.03
adj <i>according to ...</i>	ord	0.01	-0.02	0.03	0.77
	case \times ord	0.00	-0.02	0.03	0.64
	case	0.01	-0.01	0.04	0.85
vp <i>decisively supported,</i>	ord	-0.02	-0.04	0.01	0.07
	case \times ord	0.01	-0.03	0.04	0.63
	case	0.00	-0.03	0.02	0.40
pre-crit ...	ord	0.01	-0.01	0.04	0.82
	case \times ord	-0.01	-0.03	0.02	0.34
	case	-0.01	-0.03	0.02	0.33
crit <i>how,</i>	ord	0.01	-0.02	0.03	0.72
	case \times ord	0.00	-0.03	0.02	0.39
	case	-0.01	-0.04	0.02	0.28
spillover <i>so greatly ...</i>	ord	-0.01	-0.04	0.02	0.24
	case \times ord	-0.03	-0.05	0.00	0.03

Regression-path durations

At the second noun phrase, regression-path durations showed a main effect of case marking ($\hat{\beta} = 0.06$, CrI: $[0.04, 0.09]$, $\Pr(\hat{\beta} > 0) \approx 1$), as well as an interaction with word order ($\hat{\beta} = 0.04$, CrI: $[0.01, 0.07]$, $\Pr(\hat{\beta} > 0) \approx 1$), which was mainly driven by longer regression paths in

the ambiguous OVS condition compared to the ambiguous SVO condition ($\hat{\beta} = 0.06$, CrI: $[0.02, 0.10]$, $\Pr(\hat{\beta} > 0) \approx 1$). At the adjunct, regression-path durations for OVS sentences were shorter than for SVO sentences ($\hat{\beta} = -0.03$, CrI: $[-0.05, 0.00]$, $\Pr(\hat{\beta} > 0) = 0.02$). Results are shown in Table 4.17.

Table 4.17: **Experiment 2: Results for regression-path durations.** case = case marking, ord = word order.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1					
<i>A sympathizer ...</i>		See first-pass reading times			
aux	case	0.02	-0.01	0.05	0.89
<i>had</i>	ord	-0.01	-0.04	0.02	0.28
	case × ord	-0.01	-0.04	0.02	0.29
np2	case	0.06	0.04	0.09	≈ 1.00
<i>the rebels</i>	ord	0.02	-0.01	0.04	0.86
	case × ord	0.04	0.01	0.07	≈ 1.00
adj	case	0.00	-0.02	0.03	0.59
<i>according to ...</i>	ord	-0.03	-0.05	0.00	0.02
	case × ord	0.01	-0.01	0.03	0.82
vp	case	0.00	-0.03	0.04	0.52
<i>decisively supported,</i>	ord	0.00	-0.05	0.04	0.45
	case × ord	0.00	-0.04	0.05	0.55
pre-crit	case	-0.01	-0.04	0.03	0.31
...	ord	0.01	-0.02	0.04	0.79
	case × ord	0.01	-0.02	0.05	0.75
crit	case	0.01	-0.02	0.04	0.71
<i>how,</i>	ord	0.01	-0.02	0.04	0.78
	case × ord	0.00	-0.03	0.03	0.48
spillover	case	0.01	-0.02	0.04	0.75
<i>so greatly ...</i>	ord	0.01	-0.02	0.04	0.68
	case × ord	-0.02	-0.05	0.01	0.13

Total reading times

Table 4.18 lists results for total reading times. The initial noun phrase showed increased total reading times in OVS compared to SVO sentences ($\hat{\beta} = 0.03$, CrI: $[0.01, 0.06]$, $\Pr(\hat{\beta} > 0) \approx 1$). The auxiliary showed an effect of case marking, such that reading times were longer for

sentences with ambiguous case marking ($\hat{\beta} = 0.04$, CrI: [0.01,0.07], $\Pr(\hat{\beta} > 0) = 0.99$). The ambiguity effect remained in evidence on the second noun phrase ($\hat{\beta} = 0.02$, CrI: [0.00,0.05], $\Pr(\hat{\beta} > 0) = 0.96$), but the order effect was reversed, such that OVS sentences showed shorter total reading times than SVO sentences ($\hat{\beta} = -0.03$, CrI: [-0.06,0.00], $\Pr(\hat{\beta} > 0) = 0.01$).

As in first-pass reading times, the ellipsis spillover region showed an interaction between case marking and word order ($\hat{\beta} = -0.02$, CrI: [-0.05,0.00], $\Pr(\hat{\beta} > 0) = 0.05$), in this case driven by neither pair of conditions.

Table 4.18: **Experiment 2: Results for total reading times.** case = case marking, ord = word order.

region	parameter	estimate	CrI low	CrI high	$\Pr(\hat{\beta} > 0)$
np1 <i>A sympathizer ...</i>	case	0.00	-0.03	0.02	0.45
	ord	0.03	0.01	0.06	≈ 1.00
	case \times ord	0.00	-0.02	0.02	0.53
aux <i>had</i>	case	0.04	0.01	0.07	0.99
	ord	0.02	-0.01	0.05	0.91
	case \times ord	0.02	-0.01	0.04	0.90
np2 <i>the rebels</i>	case	0.02	0.00	0.05	0.96
	ord	-0.03	-0.06	0.00	0.01
	case \times ord	0.02	-0.01	0.04	0.92
adj <i>according to ...</i>	case	-0.01	-0.03	0.01	0.17
	ord	-0.02	-0.04	0.00	0.06
	case \times ord	-0.01	-0.03	0.02	0.34
vp <i>decisively supported,</i>	case	0.00	-0.02	0.03	0.61
	ord	-0.01	-0.03	0.02	0.32
	case \times ord	0.00	-0.03	0.02	0.39
pre-crit ...	case	0.01	-0.02	0.04	0.69
	ord	0.02	-0.02	0.05	0.81
	case \times ord	0.00	-0.03	0.03	0.62
crit <i>how,</i>	case	0.01	-0.01	0.03	0.78
	ord	0.02	-0.01	0.05	0.90
	case \times ord	-0.01	-0.03	0.02	0.31
spillover <i>so greatly ...</i>	case	-0.01	-0.04	0.02	0.26
	ord	0.01	-0.02	0.03	0.70
	amb \times ord	-0.02	-0.05	0.00	0.05

Reading times before and after the first fixation on the critical region

As before, we analyzed total reading times before and after the first fixation on the critical region separately in order to get a clearer picture of participants' behavior.

When only fixations made after having fixated the critical region were taken into account, there were no effects of the experimental manipulations on either rereading probabilities or rereading times when the first three regions of the antecedent – that is, the two argument noun phrases and the auxiliary – were analyzed together, nor when each region was analyzed separately.

When only fixations made prior to visiting the ellipsis site were taken into account, OVS word order led to longer total reading times on the initial noun phrase ($\hat{\beta} = 0.04$, CrI: [0.02, 0.07], $\Pr(\hat{\beta} > 0) \approx 1$), an effect that was also visible on the auxiliary ($\hat{\beta} = 0.08$, CrI: [0.05, 0.11], $\Pr(\hat{\beta} > 0) \approx 1$), but reversed on the second noun phrase ($\hat{\beta} = -0.03$, CrI: [-0.05, 0.00], $\Pr(\hat{\beta} > 0) = 0.02$).

Pre-critical total reading times on the auxiliary also showed a main effect of case marking, such that reading times were longer in the ambiguous conditions ($\hat{\beta} = 0.04$, CrI: [0.01, 0.07], $\Pr(\hat{\beta} > 0) \approx 1$), as well as an interaction indicative of a garden-path effect in the ambiguous OVS condition ($\hat{\beta} = 0.02$, CrI: [0.00, 0.05], $\Pr(\hat{\beta} > 0) = 0.95$). The second noun phrase showed the same main effect of case marking ($\hat{\beta} = 0.03$, CrI: [0.01, 0.06], $\Pr(\hat{\beta} > 0) = 0.99$) and the same interaction ($\hat{\beta} = 0.02$, CrI: [0.00, 0.05], $\Pr(\hat{\beta} > 0) = 0.95$). Results of the pre-ellipsis analysis thus closely match those of the overall analysis, with the exception of the interaction observed at the auxiliary which was not in evidence when the entire trial was taken into consideration.

Discussion

By and large, the results of Experiment 2 are similar to those of Paape (2016), though there are some important differences. The garden-path effect observed on the second noun phrase in the original self-paced reading experiment was visible at the same location in regression-path durations in the current study, suggesting that participants adopted the erroneous subject reading of the initial noun phrase in ambiguous OVS sentences and were then forced to reanalyze. Interestingly, the following adjunct region showed evidence of a speedup in regression-path durations in OVS compared to SVO sentences. A similar speedup was also observed in the original self-paced reading study, where Paape (2016) speculated that it may be due to readers trying to make up for lost time after having been slowed down by the non-canonical word order in the preceding regions.

Comprehension accuracy was comparable across the studies, even down to the changing accuracy patterns with different comprehension probes. In contrast to Paape (2016), we observed a steep drop in accuracy when the comprehension probe targeted the antecedent’s argument structure in garden-path sentences. This pattern may imply that on some trials the garden path was not resolved correctly, resulting in an incorrect representation of the antecedent. However, as the reduced accuracy did not coincide with an increase in total reading times or rereading probabilities/times for any of the relevant antecedent regions or for the ellipsis site, which might have been expected given the results from Experiment 1, we are left with the question why participants did not check their interpretations against the input more diligently, a point to which we return below.

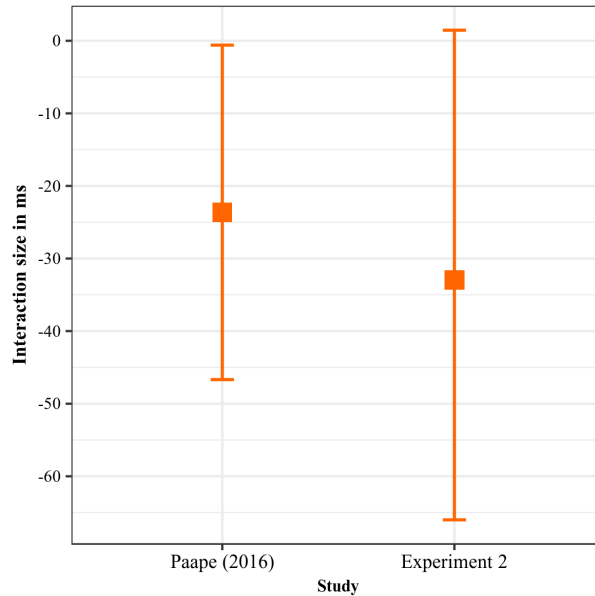
Unlike in the original study, the current experiment yielded no evidence in favor of predictive processing in the region directly before the critical region, which would have indicated that participants were expecting an elided structure. It is not clear why such an effect would appear in self-paced reading, but not in eye tracking. One possible explanation would be that self-paced reading generally puts more pressure on the processing system due to the impossibility of regressions (and possibly the lack of parafoveal preview), and that part of the computational load is shifted more towards prediction in order to minimize the impact of unexpected continuations. Furthermore, unlike eye-tracking data, self-paced reading data is composed of a series of response times with non-zero shift, due to the time it takes to perform the motor action, that is, press the space bar (Just et al., 1982; Nicenboim et al., 2016). Speculatively, additional predictive processing may be carried out during this ‘idle time’.

The crucial interaction between word order and case marking that is predicted by the reactivation hypothesis was visible in first-pass reading times at the ellipsis spillover region. Numerically, it was of the same shape as in Paape (2016); however, in the current study it was driven mainly by the ambiguous conditions. An interaction was also observed in total reading times, where it was not driven by the ambiguous conditions, just as in the original experiment.

In order to gauge the magnitude of the interaction observed in first-pass reading times at the spillover region in Experiment 2 compared to the interaction found by Paape in self-paced reading times in the third spillover region, we back-transformed both parameter estimates and their credible and confidence intervals, respectively, to the millisecond scale, as shown in Figure 4.5. We chose first-pass as opposed to total reading times as the basis for comparison because self-paced reading does not allow rereading, leading us to assume that the earliest observed effect may be the most parallel to the one observed in the original study.

As the figure shows, the parameter estimates fall between -33 ms (CrI: [1 ms, -66 ms]) in the current study and -24 ms (CI: [-1 ms, -47 ms]) in the original study. The current study

Figure 4.5: Comparison of effects between Paape (2016) and Experiment 2.



thus showed a larger estimate of the mean interaction effect, but even higher uncertainty than the original study with regard to the actual magnitude, and showed a non-zero (if low) probability of the effect being zero or positive.

Taken at face value, our result suggests that just like in the original study, having refreshed the antecedent’s memory trace through reanalysis made the trace easier to retrieve during the interpretation of the ellipsis. Yet, there are two critical discrepancies in comparison with the original study: first, no OVS mismatch effect was observed in the overall analysis, which would suggest that the amount of fit between the gap’s retrieval cue for SOV word order and the ellipsis antecedent did not influence processing at the ellipsis site. Second, as the garden-path condition showed poor comprehension accuracy, we cannot be sure that the antecedent was always retrieved correctly.

It should be noted that the reduced response accuracy in the garden-path condition, where the speedup at the critical region was observed, is not necessarily evidence against the reactivation account. Additional ‘post-interpretive processing’ may be carried out when the comprehension probe is presented (Caplan & Waters, 1999), which may involve factors beyond accessing the interpretation that was derived during reading. This type of processing may be more prone to retrieving remnants of garden-path structures than ‘true’ on-line processes (see also van Gompel, Pickering, Pearson, & Jacob, 2006).

Assuming that response accuracy *does* directly reflect parsing success, we can ask why we would observe a speedup at the ellipsis site if the antecedent was not parsed correctly, given that the findings of Experiment 1 suggest that a malformed antecedent representation leads to a slowdown at the ellipsis site.

Paape (2016) briefly considers an alternative explanation of the ambiguity advantage at the ellipsis site which rests on the assumption that discarded parses may remain active to some degree, causing comprehension errors (Christianson et al., 2001; Patson et al., 2009; Slattery et al., 2013; van Gompel et al., 2006). Under such an account, the parser would sometimes retrieve the initial, erroneous SVO representation of the first clause in the ambiguous OVS condition as the ellipsis antecedent without noticing the mistake. Processing times for the ellipsis would then be predicted to decrease, given that an SVO structure is a better match for the SOV cue from the ellipsis site than the OVS structure.

In light of the accuracy results in the current study, the possibility that participants sometimes retrieved an incorrect SVO structure in ambiguous OVS clauses cannot be discounted. Such an account, however, would raise the question why the French readers in Experiment 1 did not retrieve and integrate an incorrect structure but rather returned to the antecedent and attempted to reparse it. We take up this question in the general discussion.

In order to check whether there was any connection between processing times at the ellipsis site and probe response accuracy in Experiment 2, we performed separate analyses of first-pass reading times at the ellipsis spillover region, where the supposed reactivation effect appeared, conditional on response accuracy and probe type – that is, whether the probe targeted the assignment of thematic roles in the antecedent or not.

If the speedup was really due to erroneous retrievals, it should be most clearly visible when the subject gives an inaccurate response to a thematic role probe. This was not the case: the only subset of the data which by itself showed an interaction between word order and case marking was the one in which non-thematic role probes were responded to incorrectly, as shown by a post-hoc ancillary analysis ($\hat{\beta} = -0.20$, CrI: $[-0.33, -0.07]$, $\Pr(\hat{\beta} > 0) \approx 0$). While it should be kept in mind that this particular subset consisted of only 80 observations from 39 subjects and 17 items (as probe type was a between-items factor), the finding could be taken to imply that subjects were less able to remember meaning aspects that were *not* related to the antecedent and the ellipsis because they had focused on resolving the ellipsis correctly. In any case, we have no evidence that the garden-path effect on probe response accuracy was directly related to the speedup at the ellipsis spillover region.

In summary, the main prediction of the reactivation hypothesis was borne out in Experiment 2. However, the pattern seen in participants' overall response accuracy as well as the lack of an OVS mismatch effect at the ellipsis site cast some doubt on the result, though new problems arise when parsing failure is assumed to be responsible for the speedup, as shown by the incongruous result of the ancillary analysis.

General Discussion

We have presented two eye-tracking studies which were intended to test the predictions of the reactivation hypothesis of Paape (2016), which claims that an ellipsis antecedent becomes easier to retrieve if it has been syntactically reanalyzed at an earlier point in the sentence, due to the memory trace having received an activation boost. Experiment 1, carried out in French with three different kinds of garden-path antecedents, did not yield any evidence in favor of the reactivation hypothesis. Instead, there was some indication that, at least for one type of garden path, participants did not resolve the temporary syntactic ambiguity of the antecedent, which led to processing difficulty at the ellipsis site. Experiment 2, which was an attempt to replicate Paape's (2016) original study, yielded results that were compatible with the reactivation hypothesis; some aspects of the data, however, would allow for a different explanation of the ambiguity-induced speedup, namely that subjects may occasionally have misretrieved the initial, erroneous representation of the antecedent.

It has recently been suggested that failures to retrieve a memory target during dependency resolution may speed up processing instead of slowing it down. Nicenboim et al. (2016) found that inserting additional material between two elements of a long-distance dependency led to a slowdown in reading times for subjects with high working memory capacity, but, contrary to prediction, caused a speedup for subjects with low working memory capacity. The authors explain the finding by assuming that low-capacity readers have a higher rate of memory retrieval failure in the presence of intervening words, and that such failures are faster on average than successful retrievals. As Nicenboim et al. (2016) did not investigate garden-path structures, there was no opportunity for participants to retrieve a non-target alternative representation, as may have been the case in our Experiment 2. Nevertheless, the results call into question the implicit assumption that the time taken to process a given region is proportional to the rate of success at correctly deriving the relevant meaning.

Slattery et al. (2013) investigated small discourses like the one shown in (30), where the first sentence contains an early-closure ambiguity:

- (30) While Frank dried off the truck that was dark green was peed on by a stray dog.
Frank quickly finished drying himself off then yelled out the window at the dog.

While the results showed evidence of reanalysis in regression-path durations for the disambiguating region (*was peed on . . .*), first-pass reading times for the second clause were increased compared to a control condition, suggesting that participants perceived a clash between their initial misinterpretation (*Frank dried off the truck*) and the continuation of the discourse. This, in turn, indicates that despite participants' attempts at reanalysis, the garden-path reading still continued to interfere with processing, possibly because it was not completely inhibited.

Looking again at our Experiment 2, the situation is different from the one in (30), as the ellipsis is compatible with the correct as well as the garden-path reading of the antecedent. We would thus not expect a slowdown in processing, as there is no semantic clash, but may instead observe a speedup, because the parser has an alternative retrieval target available which yields an unlicensed but nonetheless grammatical interpretation.

In Experiment 1, a measurable effect at the ellipsis site was only observed for one stimulus type, namely for sentences with triple lexical ambiguity (TLA stimuli). For these stimuli, there was no evidence of reanalysis in the disambiguating region of the antecedent; instead, processing times were increased for the ambiguous region, suggesting competition between the different readings. Furthermore, the ambiguous condition was not easier, but rather more difficult to process at the ellipsis site, which implies that participants were having trouble completing any retrieval at all.

Taken together, the two patterns suggest that a kind of 'good-enough' processing (e.g., Karimi & Ferreira, 2016) was employed in Experiment 1, at least on some trials: despite evidence that the ambiguity was noticed by subjects, sometimes they nevertheless continued reading past the antecedent clause without resolving it, hazarding the consequences of not fully grasping the intended meaning of the whole stimulus sentence.

Note that for such a strategy to be, in fact, good enough, participants would need to have an expectation of still being able to answer the comprehension probes with a reasonable level of accuracy. Given that not all comprehension probes targeted the interpretation of the antecedent or the ellipsis site, this is not an implausible assumption. Had the wrong meaning been derived, no problem would have arisen at the ellipsis site, just as in Experiment 2, given that the ellipsis itself is compatible with both readings. However, it appears that TLA stimuli occasionally created a situation in which participants did not succeed at settling on a reading, resulting in underspecification of the antecedent's structure (Swets et al., 2008).

It would appear that underspecified structures are not valid retrieval targets, given that participants regressed to the antecedent, presumably in order to attempt a reparse.

A final remark concerns the absence of a measurable effect of antecedent reanalysis on ellipsis processing times in the remaining French sentences, that is, for subject-object inversion and reduced relative clauses. Based on comments from participants, we have argued that the inversion stimuli may sometimes have been judged as ungrammatical, which would have masked any reactivation effect, especially as the distance between the disambiguating word and the ellipsis site was much shorter than in the other stimuli. As for the reduced relative clauses, we have argued that reanalysis may not always have been carried out, presumably because the would-be garden-path structure was in fact the preferred one. However, these post-hoc hypotheses are clearly unsatisfying and need to be tested more rigorously.

To conclude, we found evidence compatible with the reactivation hypothesis only in one out of four cases of antecedent ambiguity, specifically in the same construction originally used by Paape (2016). Given the aforementioned caveats, however, we cannot be entirely certain that reactivation is indeed taking place in this case. We also found evidence suggesting that when the antecedent has a particularly difficult ambiguous structure, participants may be unable to resolve the ambiguity and end up without a retrieval target at the ellipsis site. One fruitful approach for future studies would be to probe the interpretation of the antecedent on-line, possibly in a cross-modal priming or visual world setup, so that the participants' success at parsing the garden-path structure can be assessed on a trial-by-trial basis.

Acknowledgements

The authors wish to thank Céline Pozniak for assistance during data collection and helpful comments on the French stimulus sentences. The support of Johanna Thieke, Rachel Shen, Anne Abeillé and Daniela Mertzen is also gratefully acknowledged. This work is partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program “Investissements d’Avenir” (Reference: ANR-10-LABX-0083). Experiment 2 was funded by the University of Potsdam.

CHAPTER 5

CONCLUSION

This thesis investigated possible effects of different types of antecedent complexity on the on-line processing of ellipsis. The first type of complexity studied was parsing complexity: the self-paced reading experiment presented in Chapter 2 tested whether having been confronted with a garden-path structure in the antecedent clause would lead to a measurable garden-path effect at the ellipsis site as well, assuming that the antecedent is copied into the gap as an unstructured word string (Murphy, 1985). The configuration chosen in this particular case was a well-studied subject-object ambiguity in German in which a sentence-initial noun is initially misanalyzed as a subject before reanalysis to an object-initial structure is signaled by mismatching agreement on the finite auxiliary (Hemforth, 1993; Meng & Bader, 2000). While the expected garden-path effect was observed in the antecedent clause, the ellipsis site showed no such effect, and there is thus no evidence that reparsing of the antecedent was taking place. Instead, in two of the spillover regions, an effect of antecedent-ellipsis mismatch as well as an apparent ‘reverse garden-path’ effect were observed. The former was explained by assuming that canonical antecedents match the retrieval cues set by the ellipsis site to a greater extent than non-canonical antecedents, while the latter was argued to be the result of antecedent reactivation through reanalysis within the framework of cue-based retrieval parsing (Lewis & Vasishth, 2005). There was also suggestive evidence of predictive processing of the ellipsis site that would have been consistent with earlier findings by Yoshida et al. (2013); however, a post-hoc sentence completion study did not show sluicing continuations to be preferred.

The second article (see Chapter 3) explored a different kind of antecedent complexity, namely representational complexity. While increased representational complexity goes along with increased parsing complexity, in the sense that more parsing steps are involved, no reanalysis was involved in the manipulation. Instead, antecedents were manipulated to contain additional verbs and modifiers in the complex as compared to the simple conditions. The

first self-paced reading experiment yielded no reliable evidence that increased antecedent complexity increased ellipsis processing time, as would have been predicted by (Murphy, 1985). There was also no evidence of a speedup at the ellipsis site with more complex antecedents, as would have been predicted by the complexity-based facilitation theory of Hofmeister (2011). The second self-paced reading experiment was based on the conjecture that earlier studies investigating antecedent complexity effects (Frazier & Clifton, 2000; Martin & McElree, 2008) may not have taken sufficient precautions to assure deep processing of the ellipsis, such as asking detailed comprehension questions, thereby allowing superficial processing to mask any such effects (cf. Phillips & Parker, 2014). However, the experiment failed to yield evidence of an interaction between antecedent complexity and comprehension probe type (superficial vs. detailed) at the ellipsis site. The overall pattern of results favors approaches to ellipsis processing that do not involve the creation of laborious ‘deep copies’ of antecedent material, independently of task demands, but rather ‘shallow copies’ that merely point back to the antecedent in memory. Such accounts are also straightforwardly compatible with cue-based retrieval parsing as described by Lewis and Vasishth (2005), given that antecedent chunks can be retrieved as complete units.

Chapter 4 was concerned with further testing the notion that reanalyzed antecedents may be easier to retrieve at the ellipsis site because they have received a boost in memory activation due to having been involved in additional structure-building steps. The first eye-tracking study presented in the article identified three types of garden-path construction in French, but found no evidence of a reactivation advantage for any of them. Instead, one out of the three sentence types showed evidence of processing difficulty at the ellipsis site in the form of increased total reading times. Additionally, for these sentences, an exploratory analysis of antecedent rereading after having encountered the ellipsis showed additional refixations in the ambiguous region, suggesting that readers were attempting to reparse the antecedent. Based on these findings, we conjectured that readers had, at least occasionally, engaged in ‘good enough’ processing (e.g. Karimi & Ferreira, 2016) by leaving the structure of the antecedent underspecified and ending up without a valid retrieval target. For the other two sentence types, we speculated that the ambiguities involved may have been either too difficult – in the case of subject-object inversion, which was judged as ungrammatical by some participants – or too easy – in the case of reduced relative clauses – to yield measurable reactivation effects.

The second eye-tracking study was an attempt to replicate the findings from Chapter 2, using the same materials as the original study in a different experimental paradigm. Similarly to the original study, the predicted interaction was observed in a spillover region following the critical ellipsis region, both in first-pass and total reading times. There was, however, no mismatch effect for non-canonical antecedents and no evidence of predictive

Conclusion

processing of the ellipsis. Furthermore, there was a steep drop in comprehension accuracy in the garden-path condition. Despite there being no direct evidence that this loss of accuracy was tied to the speedup at the ellipsis site, it is nevertheless possible that occasional retrievals of an incorrectly parsed antecedent may have contributed to the observed pattern.

To summarize, the evidence presented in this thesis suggests that effects of antecedent representational complexity – or ‘structural’ complexity – on ellipsis processing are either non-existent or of a very small magnitude. Parsing complexity of the antecedent, on the other hand, appears to influence ellipsis resolution only insofar as the antecedent’s processing history affects the ease with which it can be retrieved: if the correct antecedent is accessed, retrieval may be faster due to reactivation; if there are multiple antecedent representations available, an incorrect one may be retrieved, and if there is no fully analyzed antecedent available, retrieval may fail and a reparse may be attempted.

BIBLIOGRAPHY

- Ambridge, B., & Goldberg, A. E. (2008). The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics*, *19*(3), 357–389.
- Arai, M., Nakamura, C., & Mazuka, R. (2014). Predicting the unbeaten path through syntactic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 482–500.
- Arregui, A., Clifton, C., Jr., Frazier, L., & Moulton, K. (2006). Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*, *55*(2), 232–246.
- Bader, M. (2014). Das Verb *lassen* im Deutschen: Zum Zusammenspiel von Auxiliarinversion und Ersatzinfinitiv. *Zeitschrift für Sprachwissenschaft*, *33*(1), 1–44.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint, arXiv:1506.04967*.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. *1970*, 279–362.
- Bick, E. (2005). Corpuseye: Et brugervenligt web-interface for grammatisk opmærkede korpora. In P. Widell & M. Kunøe (Eds.), *10. Møde om Udforskningen af Dansk Sprog*, 7.-8. Oct., 2004 (p. 46-57).
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252.
- Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 468–481.

- Cai, Z. G., Pickering, M. J., & Sturt, P. (2013). Processing verb-phrase ellipsis in Mandarin Chinese: Evidence against the syntactic account. *Language and Cognitive Processes, 28*(6), 810–828.
- Callahan, S. M., Shapiro, L. P., & Love, T. (2010). Parallelism effects and verb activation: The sustained reactivation hypothesis. *Journal of Psycholinguistic Research, 39*(2), 101–118.
- Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(01), 77–94.
- Charvillat, A., & Kail, M. (1991). The status of ‘canonical SVO sentences’ in french: A developmental study of the on-line processing of dislocated sentences. *Journal of Child Language, 18*(03), 591–608.
- Chomsky, N. A. (1976). *Reflections on language*. New York: Pantheon.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*(4), 368–407.
- Chung, S., Ladusaw, W. A., & McCloskey, J. (1995). Sluicing and logical form. *Natural Language Semantics, 3*(3), 239–282.
- Dalrymple, M., Shieber, S. M., & Pereira, F. C. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy, 14*(4), 399–452.
- Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*(2), 193–210.
- Dickey, J. M., & Lientz, B. (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *The Annals of Mathematical Statistics, 41*(1), 214–226.
- Dickey, M. W., & Bungler, A. C. (2011). Comprehension of elided structure: Evidence from sluicing. *Language and Cognitive Processes, 26*(1), 63–78.
- Drach, E. (1937). *Grundgedanken der deutschen Satzlehre*. Frankfurt/Main: Diesterweg.
- Drummond, A. (2017). *Ibex farm* [Computer software manual]. Retrieved from <http://spellout.net/ibexfarm/>
- Dubey, A., Keller, F., & Sturt, P. (2008). A probabilistic corpus-based model of syntactic parallelism. *Cognition, 109*(3), 326–344.
- Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PLoS ONE, 8*(11), e81461.

- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164–203.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348–368.
- Ferreira, F., & Henderson, J. M. (1993). Reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 47(2), 247–275.
- Fishbein, J., & Harris, J. A. (2014). Making sense of Kafka: Structural biases induce early sense commitment for metonyms. *Journal of Memory and Language*, 76, 94–112.
- Foertsch, J., & Gernsbacher, M. A. (1994). In search of complete comprehension: Getting “minimalists” to work. *Discourse Processes*, 18(3), 271–296.
- Foraker, S., & McElree, B. (2011). Comprehension of linguistic dependencies: Speed-accuracy tradeoff evidence for direct-access retrieval from memory. *Language and Linguistics Compass*, 5(11), 764–783.
- Frazier, L. (2008). Processing ellipsis: A processing solution to the undergeneration problem. In C. B. Chang & H. J. Haynie (Eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics* (pp. 21–32). Somerville, MA.
- Frazier, L., & Clifton, C., Jr. (2000). On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research*, 29(2), 125–140.
- Frazier, L., & Clifton, C., Jr. (2001). Parsing coordinates and ellipsis: Copy α . *Syntax*, 4(1), 1–22.
- Frazier, L., & Clifton, C., Jr. (2005). The syntax-discourse divide: Processing ellipsis. *Syntax*, 8(2), 121–174.
- Frazier, L., Munn, A., & Clifton Jr., C. (2000). Processing coordinate structures. *Journal of Psycholinguistic Research*, 29(4), 343–370.
- Frazier, L., Taft, L., Roeper, T., Clifton Jr., C., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*, 12(5), 421–430.
- Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. *Advances in Psychology*, 134, 217–236.
- Fulst, S., & Phillips, C. (2004). *The source of syntactic illusions*. Poster presented at the 17th Annual CUNY Conference on Human Sentence Processing. College Park, MD.

- Gabry, J., & Goodrich, B. (2016). rstanarm: Bayesian applied regression modeling via Stan [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=rstanarm> (R package version 2.15.0-1)
- Gorrell, P. (1996). Parsing theory and phrase-order variation in German V2 clauses. *Journal of Psycholinguistic Research*, 25(1), 135–156.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114–125.
- Grinder, J., & Postal, P. M. (1971). Missing antecedents. *Linguistic Inquiry*, 2(3), 269–312.
- Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, 27(1), 135–159.
- Hankamer, J., & Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*, 7(3), 391–428.
- Hardt, D. (1993). *Verb Phrase Ellipsis: Form, Meaning, and Processing* (PhD thesis). University of Pennsylvania.
- Hartigan, J. A., & Hartigan, P. (1985). The dip test of unimodality. *The Annals of Statistics*, 13(1), 70–84.
- Hemforth, B. (1993). *Kognitives Parsing: Repräsentation und Verarbeitung sprachlichen Wissens*. Sankt Augustin: Infix.
- Hinterhölzl, R. (2006). *Scrambling, Remnant Movement, and Restructuring in West Germanic*. Oxford: Oxford University Press.
- Hirshman, E., Whelley, M. M., & Palij, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory and Language*, 28(5), 594–609.
- Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, 21(5), 1157–1164.
- Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*, 26(3), 376–405.
- Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and bayesian meta-analysis. *Journal of Memory and Language*, 94, 316–339.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.

- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228–238.
- Karimi, H., & Ferreira, F. (2016). Good-enough linguistic representations and on-line cognitive equilibrium in language processing. *The Quarterly Journal of Experimental Psychology*, *69*(5), 1013–1040.
- Kehler, A. (2000). Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, *23*(6), 533–575.
- Kennedy, C. (2003). Ellipsis and syntactic representation. In K. Schwabe & S. Winkler (Eds.), *The interfaces: Deriving and interpreting omitted structures* (pp. 29–53). Amsterdam: John Benjamins.
- Kertz, L. (2000). *Ellipsis reconsidered* (PhD thesis). University of San Diego.
- Kim, C. S., Koble, G. M., Runner, J. T., & Hale, J. T. (2011). The acceptability cline in VP ellipsis. *Syntax*, *14*(4), 318–354.
- Knoeferle, P. (2014). Conjunction meaning can modulate parallelism facilitation: Eye-tracking evidence from German clausal coordination. *Journal of Memory and Language*, *75*, 140–158.
- Kooperberg, C. (2016). logspline: Logspline density estimation routines [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=logspline> (R package version 2.1.9)
- Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*(4), 445–465.
- Lahousse, K. (2003). *The distribution of postverbal nominal subjects in French* (PhD thesis). Université de Paris 8 / Katholieke Universiteit Leuven.
- Lassotta, R., Omaki, A., & Franck, J. (2016). Developmental changes in misinterpretation of garden-path wh-questions in French. *The Quarterly Journal of Experimental Psychology*, *69*(5), 829–854.
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001.

- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.
- Logačev, P., & Vasishth, S. (2013). em2: Compute eye fixation measures, version 0.9 [Computer software manual]. Retrieved from <https://cran.r-project.org/src/contrib/Archive/em2/> (R package version 0.9)
- Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, *58*(3), 879–906.
- Martin, A. E., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1231–1239.
- McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. *Psychonomic Bulletin & Review*, *6*(3), 486–494.
- Meng, M., & Bader, M. (2000). Mode of disambiguation and garden-path strength: An investigation of subject-object ambiguities in German. *Language and Speech*, *43*(1), 43–74.
- Merchant, J. (2001). *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, J. (2013). Voice and ellipsis. *Linguistic Inquiry*, *44*(1), 77–108.
- Merchant, J. (2016). *Ellipsis: A survey of analytical approaches*. Ms., University of Chicago.
- Mitchell, D. C., Cuetos, F., Corley, M. M., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, *24*(6), 469–488.
- Montalbetti, M. M. (1984). *After binding: On the interpretation of pronouns* (PhD thesis). Massachusetts Institute of Technology.
- Müller, G. (1999). Optimality, markedness, and word order in German. *Linguistics*, *37*(5), 777–818.
- Müller, S. (2005). Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte*, *201*, 3–39.
- Murguia, E. (2004). *Syntactic identity and locality restrictions on verbal ellipsis* (PhD thesis). University of Maryland.
- Murphy, G. L. (1985). Processes of understanding anaphora. *Journal of Memory and Language*, *24*(3), 290–303.

- Murphy, G. L., & Shapiro, A. M. (1994). Forgetting of verbatim information in discourse. *Memory & Cognition*, *22*(1), 85–94.
- Nicenboim, B., Logačev, P., Gattei, C., & Vasishth, S. (2016). When high-capacity readers slow down and low-capacity readers speed up: Working memory and locality effects. *Frontiers in Psychology*, *7*(280).
- Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas – Part II. *Language and Linguistics Compass*, *10*(11), 591–613.
- Paape, D. L. J. F. (2016). Filling the silence: Reactivation, not reconstruction. *Frontiers in Psychology*, *7*(27).
- Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 280–285.
- Phillips, C., & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua*, *151*, 78–95.
- Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. *Experiments at the Interfaces*, *37*, 147–180.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110.
- Poirier, J., Wolfinger, K., Spellman, L., & Shapiro, L. P. (2010). The real-time processing of sluiced sentences. *Journal of Psycholinguistic Research*, *39*(5), 411–427.
- Potsdam, E. (2007). Malagasy sluicing and its consequences for the identity requirement on ellipsis. *Natural Language & Linguistic Theory*, *25*(3), 577–613.
- R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Retrieved from <https://www.R-project.org/> (Version 3.4.1)
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(4), 1061–1080.
- Rohde, D. (2003). *Linger*. Retrieved from <http://tedlab.mit.edu/~dr/Linger/> (Version 2.94)

- Ross, J. R. (1967). *Constraints on variables in syntax* (PhD thesis). Massachusetts Institute of Technology.
- Ross, J. R. (1969). Guess who? In R. Binnick, A. Davison, G. Green, & J. Morgan (Eds.), *Papers from the 5th Regional Meeting of the Chicago Linguistic Society* (pp. 252–286). Chicago, IL: Chicago Linguistic Society.
- Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, 2(9), 437–442.
- Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German Sentence Processing. Studies in Theoretical Psycholinguistics, Vol. 24* (pp. 65–93). Dordrecht: Springer.
- Schwartz, B., & Vikner, S. (1996/2007). The verb always leaves IP in V2 clauses. In I. Roberts (Ed.), *Comparative Grammar: Critical Concepts in Linguistics* (Vol. 5, pp. 244–302). London: Routledge.
- Simner, J., & Smyth, R. (1999). Phonological activation in anaphoric lexical access (ALA). *Brain and Language*, 68(1), 40–45.
- Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, 69(2), 104–120.
- Smith, N. J., & Levy, R. (2011). Cloze but no cigar: The complex relationship between cloze, corpus, and subjective probabilities in language processing. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1637–1642).
- Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology*, 12(3), 175–200.
- Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 47(2), 276–309.
- Stan Development Team. (2016). Stan [Computer software manual]. Retrieved from <http://mc-stan.org> (Version 2.09.0/2.15.0)
- Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, 36(1), 201–216.

- Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verbphrase anaphors. *Language and Cognitive Processes*, 5(4), 257–280.
- Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. MIT Press.
- van Gompel, R. P., Pickering, M. J., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming. *Journal of Memory and Language*, 55(3), 335–362.
- Vasisht, S., & Gelman, A. (2017). The illusion of power: How the statistical significance filter leads to overconfident expectations of replicability. *arXiv preprint arXiv:1702.00556*.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. New York: Springer.
- von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, 94, 119–133.
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cognitive Psychology*, 60(3), 158–189.
- Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E-Z Reader. *Cognition*, 111(1), 132–137.
- Wasow, T. (1972). *Anaphoric relations in English* (PhD thesis). Massachusetts Institute of Technology.
- Wellwood, A., Pancheva, R., Hacquard, V., & Phillips, C. (2017). *The anatomy of a comparative illusion*. Ms., University of Maryland.
- Williams, E. S. (1977). Discourse and logical form. *Linguistic Inquiry*, 8(1), 101–139.
- Witzel, N., Witzel, J., & Forster, K. (2012). Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research*, 41(2), 105–128.
- Xiang, M., Grove, J., & Merchant, J. (2014). *Ellipsis sites induce structural priming effects*. Ms., University of Chicago.
- Yoshida, M., Dickey, M. W., & Sturt, P. (2013). Predictive processing of syntactic structure: Sluicing and ellipsis in real-time sentence processing. *Language and Cognitive Processes*, 28(3), 272–302.

Bibliography

- Zagar, D., Pynte, J., & Rativeau IV, S. (1997). Evidence for early closure attachment on first pass reading times in French. *The Quarterly Journal of Experimental Psychology: Section A*, *50*(2), 421–438.

APPENDICES

Appendix I

Experimental materials for Chapter 2

1. a. Ein Spieler des Vereins b. Einen Spieler des Vereins c. Eine Spielerin des Vereins d. Eine Spielerin des Vereins | hatte/n | die aufdringlichen Fans | nach dem Auswärtsspiel | grob beleidigt, | aber | der Trainer | konnte | nicht | verstehen, | warum, | sodass | er | nur | enttäuscht | den Kopf schüttelte.
2. a. Ein Schüler des Schachmeisters b. Einen Schüler des Schachmeisters c. Eine Schülerin des Schachmeisters d. Eine Schülerin des Schachmeisters | hatte/n | die Schiedsrichter | während des Turniers | sehr genau beobachtet, | aber | der aufmerksame Zuschauer | fragte | sich | noch immer, | warum, | als | er | am Abend | endlich | nach Hause kam.
3. a. Ein Gönner des Künstlers b. Einen Gönner des Künstlers c. Eine Gönnerin des Künstlers d. Eine Gönnerin des Künstlers | hatte/n | die etwas seltsamen Verwandten | zu Anfang | des Mordes verdächtigt, | aber | aus den Tagebüchern | geht | nicht | hervor, | warum, | zumal | es | sich | relativ eindeutig | um Suizid handelte.
4. a. Ein Sympathisant der Opposition b. Einen Sympathisanten der Opposition c. Eine Sympathisantin der Opposition d. Eine Sympathisantin der Opposition | hatte/n | die Rebellen | laut einem Bericht | maßgeblich unterstützt, | aber | die Regierung | konnte | nicht | nachweisen, | wie, | so sehr | sich | die Untersuchungskommission | auch | bemühte.
5. a. Ein Sprecher des Pharmakonzerns b. Einen Sprecher des Pharmakonzerns c. Eine Sprecherin des Pharmakonzerns d. Eine Sprecherin des Pharmakonzerns | hatte/n | die Sportler | nach Angaben der Presse | persönlich getroffen, | aber | die Quelle | konnte | nicht | mitteilen, | wo, | sodass | die Geschichte | den meisten Lesern | wahrscheinlich | nicht sehr glaubwürdig erschien.
6. a. Ein Berater des Präsidenten b. Einen Berater des Präsidenten c. Eine Beraterin des Präsidenten d. Eine Beraterin des Präsidenten | hatte/n | die Ermittler | offensichtlich | mit Erfolg getäuscht, | aber | man | fand | nie | heraus, | wie, | denn | es | galt | nach wie vor | die höchste Geheimhaltungsstufe.
7. a. Ein Kellner des Lokals b. Einen Kellner des Lokals c. Eine Kellnerin des Lokals d. Eine Kellnerin des Lokals | hatte/n | die Stammgäste | über das geplante Skatturnier | ausgefragt, | aber | der Wirt | konnte | nicht | sagen, | warum, | da | er | offenbar | an jenem Abend | sehr beschäftigt gewesen war.
8. a. Ein Soldat der gegnerischen Streitkräfte d. Eine Soldatin der gegnerischen Streitkräfte | hatte/n | die ausgesandten Kundschafter | offenbar | in die Irre geführt, | aber | der Befehlshaber | begriff | einfach | nicht, | wie, | obwohl | ihm | die Finte | mehrmals | erklärt worden war.
9. a. Ein Gegner des umstrittenen Staudammprojekts b. Einen Gegner des umstrittenen Staudammprojekts c. Eine Gegnerin des umstrittenen Staudammprojekts d. Eine Gegnerin des umstrittenen Staudammprojekts | hatte/n | die Planer | schließlich | doch noch überzeugt, | aber | es | herrscht | Stillschweigen | darüber, | wie, | weil | niemand | sich | dem Verdacht der Bestechlichkeit | aussetzen will.
10. a. Ein Befürworter der Steuerreform b. Einen Befürworter der Steuerreform c. Eine Befürworterin der Steuerreform d. Eine Befürworterin der Steuerreform | hatte/n | die Leiter | der betroffenen Behörden | wiederholt | verbal angegriffen, | aber | es | bleibt | völlig | im Dunkeln, | weshalb, | da | das Wortgefecht | von beiden Seiten | überaus unsachlich | geführt wurde.
11. a. Ein Vertrauter des Bürgermeisters b. Einen Vertrauten des Bürgermeisters c. Eine Vertraute des Bürgermeisters d. Eine Vertraute des Bürgermeisters | hatte/n | die Ratsmitglieder | kurz vor der Wahl | auffallend häufig angerufen, | aber | heute | weiß | niemand | mehr, | warum, | wie | eine Zeitung | kürzlich | in einem Kommentar | schrieb.
12. a. Ein Sanitäter des Rettungsteams b. Einen Sanitäter des Rettungsteams c. Eine Sanitäterin des Rettungsteams d. Eine Sanitäterin des Rettungsteams | hatte/n | die Feuerwehrleute | nachdrücklich | um Hilfe

- gebeten, | aber | man | verstand | später | nicht, | warum, | bis | schließlich | Bildmaterial vom Unglücksort | das Ausmaß der Verwüstung | verständlich machte.
13. a. Ein Abgeordneter der Landtagsfraktion b. Einen Abgeordneten der Landtagsfraktion c. Eine Abgeordnete der Landtagsfraktion d. Eine Abgeordnete der Landtagsfraktion | hatte/n | die Finanzbeamten | in einem offenen Brief | gemäßregelt, | aber | fünfzig Jahre später | erscheint | es | unverständlich, | weshalb, | da | aus heutiger Sicht | wohl | kein Fehlverhalten | vorlag.
 14. a. Ein Mathematiker mit Programmierkenntnissen b. Einen Mathematiker mit Programmierkenntnissen c. Eine Mathematikerin mit Programmierkenntnissen d. Eine Mathematikerin mit Programmierkenntnissen | hatte/n | die Seitenbetreiber | über die Sicherheitslücke | informiert, | aber | der Staatsanwalt | wollte | genau | wissen, | wann, | da | dies | für den Tathergang | womöglich | äußerst entscheidend war.
 15. a. Ein Schwimmer mit zwei Beinprothesen b. Einen Schwimmer mit zwei Beinprothesen c. Eine Schwimmerin mit zwei Beinprothesen d. Eine Schwimmerin mit zwei Beinprothesen | hatte/n | die Komiteemitglieder | bezüglich der geplanten Werbekampagne | kontaktiert, | aber | es | bleibt | äußerst schleierhaft, | wann, | zumal | das Schriftstück | angeblich | zwischenzeitlich | verloren gegangen ist.
 16. a. Ein Student mit außergewöhnlichen Leistungen b. Einen Studenten mit außergewöhnlichen Leistungen c. Eine Studentin mit außergewöhnlichen Leistungen d. Eine Studentin mit außergewöhnlichen Leistungen | hatte/n | die Professoren | laut Stellungnahme des Instituts | tatkräftig unterstützt, | aber | es | war | nicht | zu erfahren, | wobei, | da | der Projektverantwortliche | nicht | für Nachfragen | zu erreichen ist.
 17. a. Ein Autor aus Bolivien b. Einen Autor aus Bolivien c. Eine Autorin aus Bolivien d. Eine Autorin aus Bolivien | hatte/n | die vier Literaturwissenschaftler | in einem 2500-Seiten-Werk | zitiert, | aber | noch | kann | niemand | sagen, | wo, | da | der Text | bislang | seltsamerweise | verschollen blieb.
 18. a. Ein Korrespondent des erfolgreichen Nachrichtensenders b. Einen Korrespondenten des erfolgreichen Nachrichtensenders c. Eine Korrespondentin des erfolgreichen Nachrichtensenders d. Eine Korrespondentin des erfolgreichen Nachrichtensenders | hatte/n | die Kollegen | vor laufender Kamera | schlechtgemacht, | aber | in einem Gespräch | konnte | nicht | festgestellt werden, | weshalb, | sodass | der Konflikt | trotz aller Entschuldigungen | ohne Zweifel | weiterhin bestehen blieb.
 19. a. Ein Violinist des Nationalorchesters b. Einen Violinisten des Nationalorchesters c. Eine Violinistin des Nationalorchesters d. Eine Violinistin des Nationalorchesters | hatte/n | die Konzertbesucher | während der halbstündigen Pause | heimlich fotografiert, | aber | der Beitrag | verriet | leider | nicht, | weshalb, | sondern | befasste | sich | eher | mit der Bildqualität.
 20. a. Ein Straßenhund mit schwarzem Fell b. Einen Straßenhund mit schwarzem Fell c. Eine Straßenhündin mit schwarzem Fell d. Eine Straßenhündin mit schwarzem Fell | hatte/n | die Kinder | bis an den Rand des Dorfes | verfolgt, | aber | niemand | konnte | sich | erklären, | weshalb, | zumal | das Tier | sich | normalerweise | vor Menschen versteckte.
 21. a. Ein Teenager ohne Schulabschluss b. Einen Teenager ohne Schulabschluss c. Eine Teenagerin ohne Schulabschluss d. Eine Teenagerin ohne Schulabschluss | hatte/n | die Talentsucher | in der Bewerbungsphase | angeschrieben, | aber | der Programmverantwortliche | fragte | sich | ernsthaft, | wozu, | denn | bemerkenswerte Fähigkeiten | wurden | an keiner Stelle | erwähnt.
 22. a. Ein Vertreter der Gewerkschaft b. Einen Vertreter der Gewerkschaft c. Eine Vertreterin der Gewerkschaft d. Eine Vertreterin der Gewerkschaft | hatte/n | die anwesenden Minister | während der Sitzung | scharf attackiert, | aber | der gesprächige Parlamentarier | wusste | selbst | nicht, | warum, | denn | er | war | nicht | dabei gewesen.
 23. a. Ein Patient mit unklaren Symptomen b. Einen Patienten mit unklaren Symptomen c. Eine Patientin mit unklaren Symptomen d. Eine Patientin mit unklaren Symptomen | hatte/n | die Krankenschwestern | dem behandelnden Arzt zufolge | mehrfach angeschrien, | aber | es | war | nicht | zu ergründen, | wieso, | obwohl | seitdem | schon | mehrere Gespräche | geführt wurden.
 24. a. Ein Biologe mit Dokortitel b. Einen Biologen mit Dokortitel c. Eine Biologin mit Dokortitel d. Eine Biologin mit Dokortitel | hatte/n | die Naturschützer | auf einer Fachkonferenz | äußerst heftig kritisiert, | aber | die anderen Teilnehmer | erinnerten | sich | nicht, | wieso, | zumal | die Diskussion | offenbar | abseits des Podiums | stattfand.
 25. a. Ein Sachverständiger aus Osteuropa b. Einen Sachverständigen aus Osteuropa c. Eine Sachverständige aus Osteuropa d. Eine Sachverständige aus Osteuropa | hatte/n | die Investoren | in der Planungsphase | eigenständig hinzugezogen, | aber | im Nachhinein | fragte | sich | so mancher Gutachter, | wieso, | da | das Ergebnis | augenscheinlich | nicht | verbessert wurde.
 26. a. Ein Redakteur der Tageszeitung b. Einen Redakteur der Tageszeitung c. Eine Redakteurin der Tageszeitung d. Eine Redakteurin der Tageszeitung | hatte/n | die maskierten Aktivisten | zu einer geheimen Videokonferenz | eingeladen, | aber | niemand | konnte | überzeugend | begründen, | wieso, | nachdem | das Vorhaben | unbeabsichtigterweise | der Öffentlichkeit | bekannt geworden war.
 27. a. Ein Spion des Inlandsgeheimdienstes b. Einen Spion des Inlandsgeheimdienstes c. Eine Spionin des Inlandsgeheimdienstes d. Eine Spionin des Inlandsgeheimdienstes | hatte/n | die Informanten | im Vorfeld

- der Verhandlungen | enttarnt, | aber | nicht einmal Experten | wussten | letztlich | zu sagen, | wie, | bis | irgendwann | eine Reinigungskraft | im Schutz der Anonymität | den entscheidenden Hinweis gab.
28. a. Ein Dolmetscher des Botschafters b. Einen Dolmetscher des Botschafters c. Eine Dolmetscherin des Botschafters d. Eine Dolmetscherin des Botschafters | hatte/n | die Gastgeber | während der Begrüßungszeremonie | empfindlich gekränkt, | aber | damals | konnte | niemand | nachvollziehen, | womit, | obwohl | die kulturellen Gepflogenheiten | der jeweils anderen Seite | auf jeden Fall | hinreichend bekannt waren.
29. a. Ein Angestellter des städtischen Verkehrsunternehmens b. Einen Angestellten des städtischen Verkehrsunternehmens c. Eine Angestellte des städtischen Verkehrsunternehmens d. Eine Angestellte des städtischen Verkehrsunternehmens | hatte/n | die Fahrgäste | mit unverschämten Äußerungen | belästigt, | aber | das Team | von Soziologen | konnte | nicht | erklären, | wieso, | sodass | der Zwischenfall | für die Wissenschaft | bis heute | rätselhaft bleibt.
30. a. Ein Aufseher des Gefängnisses b. Einen Aufseher des Gefängnisses c. Eine Aufseherin des Gefängnisses d. Eine Aufseherin des Gefängnisses | hatte/n | die verdächtigen Häftlinge | durch ein erfundenes Alibi | gedeckt, | aber | keinem der Beteiligten | war | damals | zu entlocken, | wieso, | denn | eine Aussage | hätte | wohl | gegen die Ehre verstoßen.
31. a. Ein Mitarbeiter der maroden Firma b. Einen Mitarbeiter der maroden Firma c. Eine Mitarbeiterin der maroden Firma d. Eine Mitarbeiterin der maroden Firma | hatte/n | die Geschäftsführer | in das raffinierte Veruntreuungssystem | eingeweicht, | aber | es | herrscht | Uneinigkeit | darüber, | wann, | denn | von den belastenden Dokumenten | trägt | keines | ein Datum.
32. a. Ein Geschworener des Gerichts b. Einen Geschworenen des Gerichts c. Eine Geschworene des Gerichts d. Eine Geschworene des Gerichts | hatte/n | die beiden Angeklagten | trotz richterlicher Verwarnung | direkt angesprochen, | aber | niemand im Saal | verstand | wohl | so recht, | weshalb, | bevor | die Verhandlung | überraschend | auf unbestimmte Zeit | vertagt wurde.

Appendix II

Experimental materials for Chapter 3, Experiment 1

1. Peter | sprang | über den Zaun/wollte | über den hohen Zaun | in den Nachbargarten | springen | und | sein Freund Jochen | aus der Parallelklasse | ebenfalls,/war begeistert, | weil | solche Mutproben | eigentlich | immer | Spaß machten.
2. Janine | kaufte | eine Hose/wollte | eine sommerliche Hose | für die Grillfeier | kaufen | und | ihr Begleiter Bernhard | mit den blonden Locken | ebenfalls,/wartete geduldig, | ohne | sich | Gedanken | über einen eventuell notwendigen Friseurbesuch | zu machen.
3. Der junge Physiker | entdeckte | ein neues Teilchen/wollte | durch experimentelle Untersuchungen | ein neues Elementarteilchen | entdecken | und | ein Kollege | aus Frankreich | ebenfalls,/war skeptisch, | aber | beide | ahnten | noch nichts | von der gravierenden Bedeutung | ihrer zukünftigen Erkenntnisse.
4. Der Internist | untersuchte | den eingelieferten Patienten/wollte | den eingelieferten Unfallpatienten | während der Narkose | untersuchen | und | der hinzugerufene Orthopäde | mit dem Kinnbart | ebenfalls,/wurde ohnmächtig, | da | niemand | in dem Krankenhaus | jemals | einen so übel zugerichteten Körper | gesehen hatte.
5. Das Computerprogramm | übersetzte | eine Kurzgeschichte/sollte | eine äußerst anspruchsvolle Kurzgeschichte | des gefeierten Autors | übersetzen | und | der menschliche Muttersprachler | mit dem Abschluss | in Journalistik | ebenfalls,/war ernüchert, | während | die Leser | sich | ohnehin | lieber | an das Original | hielten.
6. Der geniale Architekt | erbaute | mehrere Kirchen/durfte | mehrere eindrucksvolle Kirchen | in der Stadt | bauen | und | sein größter Konkurrent | auf diesem Gebiet | ebenfalls,/schäumte, | aber | niemand | war | zu dieser Zeit | auf die Ankunft | eines noch größeren Meisters | vorbereitet.
7. Der Professor | korrigierte | einige der Hausarbeiten/musste | einige der umfangreichen Hausarbeiten | in der Freizeit | korrigieren | und | der Tutor | mit dem zitternden Augenlid | ebenfalls,/war gespannt, | obwohl | die Noten | gar keinen | Eingang | in das Abschlusszeugnis | der Studenten | fanden.
8. Das Modell | der rumänischen Mathematikerin | erklärte | die seltsamen Lichtblitze/konnte | die seltsamen Lichtblitze | auf einfache Art | erklären | und | der Ansatz | des US-amerikanischen Teams | ebenfalls,/war vielversprechend, | doch | das Rätsel | der fremdartigen Geräusche | wurde | erst Jahrzehnte später | von einem Militärpiloten | gelöst.
9. Eine Gruppe | von Geologen | erforschte | den Vulkankrater/sollte | den gigantischen Vulkankrater | bis ins letzte Detail | erforschen | und | ein ferngesteuerter Roboter | aus einem neuartigen Material |

- ebenfalls,/verrostete derweil, | bis | dem Institut | schließlich | das Geld | ausging | und | die Finanzierung | gestrichen wurde.
10. Der erfolgreiche Autohersteller | präsentierte | einen robusten Geländewagen/konnte | auf der Fachmesse | einen robusten Geländewagen | präsentieren | und | der ambitionierte Technikkonzern | mit der riesigen Forschungsabteilung | ebenfalls,/war neidisch, | denn | Marktanalysen | sagten | für dieses Segment | ein rasantes Wachstum | voraus.
 11. Die Lasertherapie | entfernte | zuverlässig | Tumoren/konnte | Tumoren | bei Menschen | zuverlässig | entfernen | und | der chirurgische Eingriff | mit anschließender Chemotherapie | ebenfalls,/wurde unbeliebter, | sodass | das Leben | vieler Patienten | nachweislich | deutlich | verlängert werden konnte.
 12. Der kleine Küstenstaat | kündigte | alle internationalen Verträge/wollte | alle internationalen Verträge | noch vor Jahresende | kündigen | und | der größte Erdölexporteur | in der Region | ebenfalls,/lieferte nicht, | obwohl | nicht einmal | die herrschende Elite | daraus | einen Vorteil zog.
 13. Der gute alte Deich | schützte | die Kleinstadt/sollte | die Kleinstadt | vor den erwarteten Wassermassen | schützen | und | der absenkbarer Schutzwall | einer japanischen Firma | ebenfalls,/wurde abgerissen, | doch | im Jahre 2010 | gab | es | ein Erdbeben | und | der Ort | wurde | überflutet.
 14. Der betagte Alpinist | bestieg | einen besonders tückischen Gipfel/wollte | einen besonders tückischen Gipfel | mit minimaler Ausrüstung | besteigen | und | sein Herausforderer | mit dem auffälligen Haarschnitt | ebenfalls,/war verblüfft, | während | die Sponsoren | sich | über ihren genialen Coup | freuten.
 15. Der Auftragsmörder | vernichtete | alle Beweise/konnte | alle Beweise | vor der Durchsuchung | vernichten | und | sein Komplize | bei der Polizei | ebenfalls,/erhängte sich, | jedoch | ließ | sich | der Tathergang | anhand von Indizien | trotzdem | relativ genau | rekonstruieren.
 16. Der Hausmeister | des Raubkatzengebäudes | reinigte | die Käfige | zu selten/sollte | die Käfige | laut Vertrag | eigenhändig | reinigen | und | der Elefantenwärter | mit dem Raucherhusten | ebenfalls,/war krankgeschrieben, | sodass | der Maßstab | für Sauberkeit | im städtischen Zoo | nicht | allzu hoch war.
 17. Ein Gesandter | der Bundesregierung | reiste | in die Krisenregion/wollte | in die Krisenregion | jenseits der Sicherheitszone | reisen | und | ein Journalist | von einer namhaften Tageszeitung | ebenfalls,/wird vermisst, | während | weiterhin | täglich | ganze Dörfer | in Schutt und Asche | gelegt wurden.
 18. Die Armee | räumte | einige wichtige Feldlager/ließ | nach dem Gefecht | einige wichtige Feldlager | räumen | und | der kluge Befehlshaber | der Aufständischen | ebenfalls,/rückte vor, | ohne | allerdings | die Sicherung | seiner Versorgungslinien | zu vernachlässigen.
 19. Der Verfassungsschutz | beobachtete | weiterhin | mehrere Parteifunktionäre/ließ | weiterhin | ohne jede gesetzliche Grundlage | mehrere prominente Parteifunktionäre | beobachten | und | ein neuer Zweig | des Bundeskriminalamts | ebenfalls,/wurde gegründet, | anstatt | dass | man | über die Zweckmäßigkeit | der getroffenen Maßnahmen | diskutiert hätte.
 20. Karl-Heinz | fütterte | die Kühe/musste | nach dem Mittagessen | die Kühe | füttern | und | sein Nachbar Thorben | vom alten Gutshof | ebenfalls,/pflügte, | so | wie | es | seit zwanzig Jahren | jeden Tag | gewesen war.
 21. Der außerirdische Organismus | reagierte | auf Radioaktivität/schien | auf Radioaktivität | durch schnellere Zellteilung | zu reagieren | und | das nachgezüchtete Gewebe | aus dem Labor | ebenfalls,/wurde zerstört, | sofern | man | den Angaben | der ungewohnt geheimnistuerischen Raumfahrtbehörde | in diesem Fall | trauen konnte.
 22. Der Onlineversand | änderte | die Geschäftsbedingungen/musste | die Geschäftsbedingungen | zugunsten der Kunden | ändern | und | der Internet-Bezahldienst | als wichtigster Partner | war pleite, | denn | zu viele Käufer | hatten | sich | aus Ärger | über die Rückvergütungsregelung | abgewendet.
 23. Der amtierende Weltmeister | fuhr | neben der Strecke/musste | wegen der Trümmerteile | neben der Strecke | fahren | und | sein Teamkollege | auf dem zweiten Ranglistenplatz | ebenfalls,/schied aus, | sodass | nach dem Rennen | tatsächlich | ein wenig Bewegung | in die Punktetabelle | kam.
 24. Das Gesundheitsamt | maß | erhöhte Bleikonzentrationen/hatte | erhöhte Bleikonzentrationen | im Wasser | festgestellt | und | der Naturschutzbund | mit seinen schärferen Richtlinien | ebenfalls,/war alarmiert, | nachdem | Anwohner | von großen Mengen | toter Fische | in dem See | berichtet hatten.
 25. Der Raumschiffkapitän | bediente | vielerlei hochsensible Elektronik/musste | auf der Brücke | vielerlei hochsensible Elektronik | bedienen | und | der Navigator | vom Planeten Kepler-186f | ebenfalls,/war beeindruckt, | aber | keine Technologie | im Universum | hätte | die Katastrophe | verhindern können.
 26. Der belagerte König | ernannte | einen Unterhändler/ließ | angesichts der verfahrenen Lage | einen Unterhändler | ernennen | und | der Thronfolger | in seinem heruntergekommenen Kommandozelt | ebenfalls,/frohlockte, | aber | der Krieg | sollte | sich | trotzdem | noch über Jahre | hinziehen.
 27. Der russische Wissenschaftler | dressierte | mehrere Minensuchpinguine/wollte | in einem Pilotprojekt | mehrere Minensuchpinguine | dressieren | und | der internationale Forschungsfonds | als Kooper-

ationspartner | ebenfalls,/war ahnungslos, | was | nach Bekanntwerden | heftige Kritik | vonseiten verschiedener Tierschutzorganisationen | provozierte.

28. Der Verwaltungsangestellte | installierte | das Update/sollte | das Update | für die Datenbankanwendung | installieren | und | die Sekretärin | im Büro nebenan | ebenfalls,/telefonierte gerade, | als | die IT-Abteilung | plötzlich | im gesamten Gebäude | den Strom | abschaltete.

Appendix III

Experimental materials for Chapter 3, Experiment 2

1. The history professor | understood | ancient Tibetan mythology,/ancient Tibetan mythology's artistic and intellectual appeal, | but | later | it became clear | that | the bleary-eyed and apathetic graduate students | did not, | although | the stories about demonic zombies | stuck | with them.
2. The fashion designer | did not like | the satin blouse,/the satin blouse's revealing and provocative cut, | but | he | saw | that | the bossy supermodel from the Caucasus | did, | while | insisting | that | the cashmere scarf | was absolutely loathsome.
3. The private detective | lamented | the murder investigation,/the newest murder investigation's possible political consequences, | but | it was obvious | that | the arrogant district attorney | quite naturally | did not, | as | high-profile cases | served to | increase | his renown.
4. The head physician | appreciated | the accurate diagnosis,/the carefully written diagnosis' accuracy and thoroughness, | but | it was tragic | that | the suffering patient and his family | did not, | as | there was | not much room left | for hope.
5. The defense player | denied | the outrageous accusation,/the young assistant referee's completely outrageous accusation, | but | it was unfortunate | that | the uncertain team manager | rather understandably | did not, | given | that | the alleged perpetrator | was known for | foul play.
6. The police detective | did not fear | the brutal robber,/the brutal robber's entirely premature parole release, | but | he | accepted | that | clearly and understandably | the traumatized victim | did, | even though | tight post-release supervision | had been ordered.
7. The hard-working employees | enjoyed | the long weekend,/the long weekend's vital and well-deserved respite, | but | everybody | understood | that | the recently divorced department chief | undoubtedly | did not, | because | working | took | his mind | off his personal problems.
8. The impressionable girl | relished | the comedic play,/the famous author's recently released comedic play, | but | it was plain | that | of course | the jaundiced newspaper critic | did not, | speculating | that | the writer | might be | mentally deranged.
9. The elderly lady | pitied | the stray dog,/the ill-fated stray dog's little newborn puppies, | but | it disturbed her | that | the intimidating and cold-hearted veterinarian | obviously | did not, | as | he | suggested | putting it to sleep | with an eerie gleam | in his eyes.
10. The gourmet chef | praised | the low-calorie pastry,/the low-calorie vegetarian pastry's rather unconventional taste, | but | he | noticed | that | the less sophisticated diabetic | quite unsurprisingly | did not, | even though | he admitted | the cauliflower | was a nice touch.
11. The advanced students | loved | the afternoon session,/the late afternoon session's many illustrative examples, | but | as of late | it was evident | that | the mathematics lecturer | did not, | as | the time-consuming preparation | really | exhausted her.
12. The cheerless bishop | did not enjoy | the evening ball,/the evening ball's excessive pomp and jollity, | but | one could see | that | without a doubt | the blossoming queen | did, | basking | in the admiration | of her subjects.
13. The sleepless groom | dreaded | the upcoming wedding,/the upcoming wedding's excessively fancy dinner reception, | but | the chaplain | was delighted to see | that | the charming bride | in her happiness | did not, | even though | the revealing dress | made her | a little bit nervous.
14. The grouchy dieter | hated | the fitness plan,/the fitness plan's difficult and tiring exercises, | but | he | discovered | that | the very enthusiastic personal trainer | clearly | did not, | remarking | that | several celebrities | exercised | according to his workout regimen.
15. The successful businesswoman | accepted | the unplanned pregnancy,/the young secretary's unplanned and inconvenient pregnancy, | but | she | was told | that | her reluctant boyfriend of five years | did not, | knowing | that | he | was | severely lacking in fatherly qualities.

Appendices

16. The thrill-seeking surfer | did not mind | the cold water,/the secluded bay's crystal-clear and ice-cold water, | but | he | could tell | that | the hypersensitive sunbather on the beach | did, | insisting | that | too much salt | would | surely | ruin his tan.
17. The shrewd investor | did not fear | the stock prices,/the overblown stock prices' nearly inevitable decline, | but | it was common knowledge | that | the apprehensive broker and his assistant | did, | as | they | had neglected to | hedge | the risks.
18. The rescue workers | awaited | the relief forces,/the local military's competent disaster relief forces, | but | citizens | reported | that | the stout-hearted fireman with the cigar | did not, | starting | to dig | through the rubble | with the help of his colleagues.
19. The intrepid adventurer | savored | the hunting trip,/the hunting trip's savage and intoxicating thrill, | but | it became clear | that | the archaeologist from some obscure university | did not, | and | thus | preferred to | stay behind | at the campsite.
20. The traveling salesman | did not fancy | small commuter planes,/small commuter planes' engine noise and incessant tottering, | but | he | was aware | that | the fledgling pilot's little son | undeniably | did, | always | looking forward to | being allowed on board.
21. The gullible newspaperman | did not distrust | the food company,/the food company's aggravatingly shiftily media spokesperson, | but | it was reasonable | that | the crusading nutritionist from the capital | did, | knowing | that | profit | was | always | more important than consumer interests.
22. The hired crew | did not respect | the old captain,/the old captain's stubbornness and physical tenacity, | but | the diary entries | showed | that | the luckless whaler's sullen first mate | did, | partly | because | the man | could | really | hold his liquor.
23. The exhausted nurse | welcomed | the afternoon break,/the afternoon break's aimlessly floating small talk, | but | everyone | felt | that | the highly spirited medical student | usually | did not, | impatiently | watching | the hands of the clock | advance.
24. The misanthropic poet | did not ponder | the publication offer,/the recently founded agency's generous publication offer, | but | it was obvious | that | the utterly destitute science fiction writer | did, | wondering | who | had sent in | his manuscript.
25. The seasoned investigator | dreaded | the scheduled meeting,/the scheduled meeting's entirely fruitless legal debates, | but | soon | he | realized | that | the dashing attractive top litigation attorney | did not, | as | she | seemed | to enjoy | these kinds of things.
26. The enterprising chairman | did not second-guess | the generous proposal,/the shady insurance company's generous collaboration proposal, | but | what was crucial | was | that | the staunch legal expert from Compliance | did, | insisting | that | the risk of judicial complications | was too high.
27. The conscientious columnist | did not appreciate | the weekend gossip,/the weekend gossip's insolent tactlessness and malice, | but | he | had to | take into account | that | the money-hungry newspaper owner | most certainly | did, | as | it | tended | to boost sales.
28. The French babysitter | adored | the identical twins,/the identical twins' matching green wool hats, | but | it turned out | that | the annoyed home economics teacher | absolutely | did not, | because | it was impossible | to tell | the two of them | apart.
29. The frustrated wife | did not regret | the marital quarrel,/the marital quarrel's hurtful and aggressive tone, | but | it was interesting | that | to his own surprise | the husband | did, | perhaps | because | love | was | stronger than pride | after all.
30. The recalcitrant teenager | despised | the foster parents,/the new classmate's boastfully rich foster parents, | but | he | was furious | that | the little preschooler | in his callowness | did not, | because | they | gave him | candy | sometimes.
31. The ill-natured grandmother | did not cherish | the extended family,/the extended family's by now traditional get-togethers, | but | it became clear | that | remarkably | the shy and melancholy youngster | did, | because | everyone | accepted | his personality.
32. The dutiful soldier | did not recognize | the battalion commander,/the battalion commander's icy and ruthless gaze, | but | it was his undoing | that | the freshly apprehended deserter | most definitely | did, | as | he | had first-handedly witnessed | the massacre.
33. The socialist mayor | did not condemn | the violent protest,/the unpaid coal miners' increasingly violent protest, | but | it proved disastrous | that | the city's hard-line chief of police | did, | and | haphazardly | decided | to meet force with force.
34. The ambitious researcher | did not need | the impressive grant,/the well-known private foundation's rather impressive grant, | but | she was informed | that | apparently | the multinational corporation's staff scientists | did, | given | that | various divisions | had applied | for it.

35. The egocentric actress | did not desire | the prestigious award,/the film society's prestigious lifetime achievement award, | but | she | observed | that | the notoriously tight-lipped star director | evidently | did, | shifting | nervously | in his seat | during the ceremony.
36. The department head | did not value | the brilliant engineer,/the brilliant propulsion engineer's honest professional opinion, | but | fortunately | it was known | that | the car company's new CEO | did, | because | the firm's success | depended on | people like him.

Appendix IV

Experimental materials for Chapter 4, Experiment 1

Note that for Experiment 2, the same materials as in Chapter 2 were used.

RRC stimuli

1.
 - a. Le politicien | contredit | durant le débat | a proposé | des mesures | à réaliser | en vitesse, | mais tous les quotidiens | de tendance libérale | se demandaient | pourquoi, | vu que | ces derniers temps | il n'y a pas eu | de menaces terroristes | à prendre au sérieux.
 - b. La politicienne | contredite | durant le débat | a proposé | des mesures | à réaliser | en vitesse, | mais tous les quotidiens | de tendance libérale | se demandaient | pourquoi, | vu que | ces derniers temps | il n'y a pas eu | de menaces terroristes | à prendre au sérieux.
 - c. Le politicien | contredit | durant le débat | a proposé | des mesures | à réaliser | en vitesse, | mais tous les quotidiens | de tendance libérale | les critiquaient, | vu que | ces derniers temps | il n'y a pas eu | de menaces terroristes | à prendre au sérieux.
 - d. La politicienne | contredite | durant le débat | a proposé | des mesures | à réaliser | en vitesse, | mais tous les quotidiens | de tendance libérale | les critiquaient, | vu que | ces derniers temps | il n'y a pas eu | de menaces terroristes | à prendre au sérieux.
2.
 - a. Le héros | décrit | dans cette histoire | a vaincu | une bête féroce, | mais aucun conte | du romancier | ne mentionne | comment, | peut-être | parce qu'il | n'était pas | présent à l'époque.
 - b. Les héros | décrits | dans cette histoire | ont vaincu | une bête féroce, | mais aucun conte | du romancier | ne mentionne | comment, | peut-être | parce qu'il | n'était pas | présent à l'époque.
 - c. Le héros | décrit | dans cette histoire | a vaincu | une bête féroce, | mais aucun conte | du romancier | ne le mentionne, | peut-être | parce qu'il | n'était pas | présent à l'époque.
 - d. Les héros | décrits | dans cette histoire | ont vaincu | une bête féroce, | mais aucun conte | du romancier | ne le mentionne, | peut-être | parce qu'il | n'était pas | présent à l'époque.
3.
 - a. Le navire | détruit | pendant la guerre | avait rejoint | le port, | mais le professeur | d'histoire | ne pouvait pas dire | quand, | laissant | un peu désillusionnée | la jeune collègue | durant leur rendez-vous.
 - b. Les navires | détruits | pendant la guerre | avaient rejoint | le port, | mais le professeur | d'histoire | ne pouvait pas dire | quand, | laissant | un peu désillusionnée | la jeune collègue | durant leur rendez-vous.
 - c. Le navire | détruit | pendant la guerre | avait rejoint | le port, | mais le professeur | d'histoire | n'en savait rien, | laissant | un peu désillusionnée | la jeune collègue | durant leur rendez-vous.
 - d. Les navires | détruits | pendant la guerre | avaient rejoint | le port, | mais le professeur | d'histoire | n'en savait rien, | laissant | un peu désillusionnée | la jeune collègue | durant leur rendez-vous.
4.
 - a. L'étudiant | distrait | pendant la classe | était parti, | mais même le directeur | de l'école | ne savait pas | pourquoi, | puisqu'il | n'y avait pas | de lettre | ni d'avertissement.
 - b. L'étudiante | distraite | pendant la classe | était partie, | mais même le directeur | de l'école | ne savait pas | pourquoi, | puisqu'il | n'y avait pas | de lettre | ni d'avertissement.
 - c. L'étudiant | distrait | pendant la classe | était parti, | mais même le directeur | de l'école | ne le savait pas, | puisqu'il | n'y avait pas | de lettre | ni d'avertissement.
 - d. L'étudiante | distraite | pendant la classe | était partie, | mais même le directeur | de l'école | ne le savait pas, | puisqu'il | n'y avait pas | de lettre | ni d'avertissement.
5.
 - a. Le dernier chapitre | soustrait | du livre | avait | l'air ingénieux, | mais effectivement | personne | ne pouvait dire | exactement | pourquoi, | d'autant que | beaucoup d'intellectuels | n'avaient même pas | compris l'introduction.

- b. Les derniers chapitres | soustraits | du livre | avaient | l'air ingénieux, | mais effectivement | personne | ne pouvait dire | exactement | pourquoi, | d'autant que | beaucoup d'intellectuels | n'avaient même pas | compris l'introduction.
- c. Le dernier chapitre | soustrait | du livre | avait | l'air ingénieux, | mais effectivement | personne | ne le savait, | d'autant que | beaucoup d'intellectuels | n'avaient même pas | compris l'introduction.
- d. Les derniers chapitres | soustraits | du livre | avaient | l'air ingénieux, | mais effectivement | personne | ne le savait, | d'autant que | beaucoup d'intellectuels | n'avaient même pas | compris l'introduction.
6. a. L'enseignant | inscrit | au cours de mathématiques avancées | s'était tué, | mais le porte-parole | de la police | n'avait pas dit | comment, | de sorte que | les villageois | au café | ne pouvaient que spéculer.
- b. L'enseignante | inscrite | au cours de mathématiques avancées | s'était tuée, | mais le porte-parole | de la police | n'avait pas dit | comment, | de sorte que | les villageois | au café | ne pouvaient que spéculer.
- c. L'enseignant | inscrit | au cours de mathématiques avancées | s'était tué, | mais le porte-parole | de la police | n'en avait | rien révélé, | de sorte que | les villageois | au café | ne pouvaient que spéculer.
- d. L'enseignante | inscrite | au cours de mathématiques avancées | s'était tuée, | mais le porte-parole | de la police | n'en avait | rien révélé, | de sorte que | les villageois | au café | ne pouvaient que spéculer.
7. a. Le secrétaire | instruit | de la défaite | aurait été trouvé | dans un bar, | mais sur la base du rapport militaire | il est impossible | de dire | par qui, | d'autant que | les passages pertinents | ont été noircis.
- b. La secrétaire | instruite | de la défaite | aurait été trouvée | dans un bar, | mais sur la base du rapport militaire | il est impossible | de dire | par qui, | d'autant que | les passages pertinents | ont été noircis.
- c. Le secrétaire | instruit | de la défaite | aurait été trouvé | dans un bar, | mais sur la base du rapport militaire | il est impossible | de le vérifier, | d'autant que | les passages pertinents | ont été noircis.
- d. La secrétaire | instruite | de la défaite | aurait été trouvée | dans un bar, | mais sur la base du rapport militaire | il est impossible | de le vérifier, | d'autant que | les passages pertinents | ont été noircis.
8. a. Le fonctionnaire | interdit | de séance | a reçu | une convocation, | mais aucun de ses collègues | ne pouvait encore | expliquer | pourquoi, | jusqu'à ce que | la police secrète | l'arrête | pour haute trahison.
- b. La fonctionnaire | interdite | de séance | a reçu | une convocation, | mais aucun de ses collègues | ne pouvait encore | expliquer | pourquoi, | jusqu'à ce que | la police secrète | l'arrête | pour haute trahison.
- c. Le fonctionnaire | interdit | de séance | a reçu | une convocation, | mais aucun de ses collègues | ne pouvait encore | l'expliquer, | jusqu'à ce que | la police secrète | l'arrête | pour haute trahison.
- d. La fonctionnaire | interdite | de séance | a reçu | une convocation, | mais aucun de ses collègues | ne pouvait encore | l'expliquer, | jusqu'à ce que | la police secrète | l'arrête | pour haute trahison.
9. a. Le fruit exotique | introduit | au village | était | fortement toxique, | mais sur le moment | les scientifiques de l'université | n'avaient pas | découvert | pourquoi, | alors même que | le gouvernement | leur avait fourni | des moyens financiers considérables.
- b. Les fruits exotiques | introduits | au village | étaient | fortement toxiques, | mais sur le moment | les scientifiques de l'université | n'avaient pas | découvert | pourquoi, | alors même que | le gouvernement | leur avait fourni | des moyens financiers considérables.
- c. Le fruit exotique | introduit | au village | était | fortement toxique, | mais sur le moment | les scientifiques de l'université | ne l'avaient pas | découvert, | alors même que | le gouvernement | leur avait fourni | des moyens financiers considérables.
- d. Les fruits exotiques | introduits | au village | étaient | fortement toxiques, | mais sur le moment | les scientifiques de l'université | ne l'avaient pas | découvert, | alors même que | le gouvernement | leur avait fourni | des moyens financiers considérables.
10. a. L'artiste | peint | dans son atelier | avait rencontré | une comtesse, | mais le guide avait malheureusement oublié | de vérifier | où, | de sorte que | les touristes | se sont plaints | à l'agence de voyage.
- b. L'artiste | peinte | dans son atelier | avait rencontré | une comtesse, | mais le guide avait malheureusement oublié | de vérifier | où, | de sorte que | les touristes | se sont plaints | à l'agence de voyage.
- c. L'artiste | peint | dans son atelier | avait rencontré | une comtesse, | mais le guide avait malheureusement oublié | d'en vérifier | le lieu, | de sorte que | les touristes | se sont plaints | à l'agence de voyage.
- d. L'artiste | peinte | dans son atelier | avait rencontré | une comtesse, | mais le guide avait malheureusement oublié | d'en vérifier | le lieu, | de sorte que | les touristes | se sont plaints | à l'agence de voyage.
11. a. Le roi éternel | prédit | dans les écritures | aurait accompli | un miracle, | mais de nos jours il est difficile | de dire | quand, | en particulier | parce que | la Curie refuse | de rendre accessibles | ses archives.
- b. La reine éternelle | prédite | dans les écritures | aurait accompli | un miracle, | mais de nos jours il est difficile | de dire | quand, | en particulier | parce que | la Curie refuse | de rendre accessibles | ses archives.

- c. Le roi éternel | prédit | dans les écritures | aurait accompli | un miracle, | mais de nos jours il est difficile | d'en découvrir | la nature, | en particulier | parce que | la Curie refuse | de rendre accessibles | ses archives.
- d. La reine éternelle | prédite | dans les écritures | aurait accompli | un miracle, | mais de nos jours il est difficile | d'en découvrir | la nature, | en particulier | parce que | la Curie refuse | de rendre accessibles | ses archives.
12. a. Le robot | produit | à l'usine | doit traiter | des pièces uniques, | mais | pour le chef d'équipe c'est un vrai problème | de décider | quand, | vu qu'il y a | plein d'autres | commandes importantes | à finir.
- b. Les robots | produits | à l'usine | doivent traiter | des pièces uniques, | mais | pour le chef d'équipe c'est un vrai problème | de décider | quand, | vu qu'il y a | plein d'autres | commandes importantes | à finir.
- c. Le robot | produit | à l'usine | doit traiter | des pièces uniques, | mais | pour le chef d'équipe c'est un vrai problème | d'en trouver | le temps, | vu qu'il y a | plein d'autres | commandes importantes | à finir.
- d. Les robots | produits | à l'usine | doivent traiter | des pièces uniques, | mais | pour le chef d'équipe c'est un vrai problème | d'en trouver | le temps, | vu qu'il y a | plein d'autres | commandes importantes | à finir.
13. a. Le texte | reconstruit | à partir d'anciennes épigraphes | contient | des contradictions, | mais pour les philosophes des temps modernes il est difficile | d'établir | lesquelles, | car les penseurs | de la Renaissance | ont un style | trop opaque.
- b. Les textes | reconstruits | à partir d'anciennes épigraphes | contiennent | des contradictions, | mais pour les philosophes des temps modernes il est difficile | d'établir | lesquelles, | car les penseurs | de la Renaissance | ont un style | trop opaque.
- c. Le texte | reconstruit | à partir d'anciennes épigraphes | contient | des contradictions, | mais pour les philosophes des temps modernes il est | difficile | de les établir, | car les penseurs | de la Renaissance | ont un style | trop opaque.
- d. Les textes | reconstruits | à partir d'anciennes épigraphes | contiennent | des contradictions, | mais pour les philosophes des temps modernes il est | difficile | de les établir, | car les penseurs | de la Renaissance | ont un style | trop opaque.
14. a. Le sultan | satisfait | depuis peu | avait | un grand harem, | mais les visiteurs | du palais | ne savaient pas | où, | et n'avaient | même pas | vraiment | le droit de chercher.
- b. Les sultans | satisfaits | depuis peu | avaient | un grand harem, | mais les visiteurs | du palais | ne savaient pas | où, | et n'avaient | même pas | vraiment | le droit de chercher.
- c. Le sultan | satisfait | depuis peu | avait | un grand harem, | mais les visiteurs | du palais | ne savaient pas | en trouver | l'emplacement, | et n'avaient | même pas | vraiment | le droit de chercher.
- d. Les sultans | satisfaits | depuis peu | avaient | un grand harem, | mais les visiteurs | du palais | ne savaient pas | en trouver | l'emplacement, | et n'avaient | même pas | vraiment | le droit de chercher.
15. a. L'acteur | séduit | à la soirée | a rencontré | une journaliste, | mais les données | prises du portable | ne révèlent pas | où, | bien que | les enquêteurs | aient fait appel | aux techniciens spécialisés.
- b. L'actrice | séduite | à la soirée | a rencontré | une journaliste, | mais les données | prises du portable | ne révèlent pas | où, | bien que | les enquêteurs | aient fait appel | aux techniciens spécialisés.
- c. L'acteur | séduit | à la soirée | a rencontré | une journaliste, | mais les données | prises du portable | n'en révèlent pas | l'endroit, | bien que | les enquêteurs | aient fait appel | aux techniciens spécialisés.
- d. L'actrice | séduite | à la soirée | a rencontré | une journaliste, | mais les données | prises du portable | n'en révèlent pas | l'endroit, | bien que | les enquêteurs | aient fait appel | aux techniciens spécialisés.
16. a. Le dictionnaire | traduit | de l'anglais | contenait | des familles de mots | en plus, | mais la table | des matières | n'indique pas | lesquelles, | de façon | qu'il serait | extrêmement laborieux | de les identifier.
- b. Les dictionnaires | traduits | de l'anglais | contenaient | des familles de mots | en plus, | mais la table | des matières | n'indique pas | lesquelles, | de façon | qu'il serait | extrêmement laborieux | de les identifier.
- c. Le dictionnaire | traduit | de l'anglais | contenait | des familles de mots | en plus, | mais la table | des matières | ne les indique pas, | de façon | qu'il serait | extrêmement laborieux | de les identifier.
- d. Les dictionnaires | traduits | de l'anglais | contenaient | des familles de mots | en plus, | mais la table | des matières | ne les indique pas, | de façon | qu'il serait | extrêmement laborieux | de les identifier.

TLA stimuli

1. C Les adolescents ont entendu qu'il y a un pianiste handicapé formidable sur YouTube.

- a. La vidéo | illustre | les touche, | mais il s'avère | curieusement | difficile | de formuler | pourquoi, | sauf que | le triomphe | de la créativité | sur la destinée | est exaltant à regarder.
 - b. Les vidéos | illustres | les touchent, | mais il s'avère | curieusement | difficile | de formuler | pourquoi, | sauf que | le triomphe | de la créativité | sur la destinée | est exaltant à regarder.
 - c. La vidéo | illustre | les touche, | mais il s'avère | curieusement | difficile | d'en formuler | la raison, | sauf que | le triomphe | de la créativité | sur la destinée | est exaltant à regarder.
 - d. Les vidéos | illustres | les touchent, | mais il s'avère | curieusement | difficile | d'en formuler | la raison, | sauf que | le triomphe | de la créativité | sur la destinée | est exaltant à regarder.
2. C Pour cacher les trous, les ouvriers du chantier attachent de grandes bâches au bateau.
- a. Le montage | fixe | les voile, | mais le patron | sait | qu'on | ne peut | garantir | pour combien de temps, | même si | la visite | d'acheteurs potentiels | se déroule | sans problème.
 - b. Les montages | fixes | les voilent, | mais le patron | sait | qu'on | ne peut | garantir | pour combien de temps, | même si | la visite | d'acheteurs potentiels | se déroule | sans problème.
 - c. Le montage | fixe | les voile, | mais le patron | sait | qu'on | ne peut | le garantir, | même si | la visite | d'acheteurs potentiels | se déroule | sans problème.
 - d. Les montages | fixes | les voilent, | mais le patron | sait | qu'on | ne peut | le garantir, | même si | la visite | d'acheteurs potentiels | se déroule | sans problème.
3. C La directrice du groupe découvre que quelques collaborateurs ont détourné d'importantes sommes d'argent.
- a. Le projet | informe | les aide, | mais grâce à | la mauvaise comptabilité, | il n'y a | aucun moyen | de saisir | depuis quand, | incitant | une investigation interne | par | des examinateurs minutieux.
 - b. Les projets | informes | les aident,, | mais grâce à | la mauvaise comptabilité, | il n'y a | aucun moyen | de saisir | depuis quand, | incitant | une investigation interne | par | des examinateurs minutieux.
 - c. Le projet | informe | les aide, | mais grâce à | la mauvaise comptabilité, | il n'y a | aucun moyen | d'en saisir | les coupables, | incitant | une investigation interne | par | des examinateurs minutieux.
 - d. Les projets | informes | les aident, | mais grâce à | la mauvaise comptabilité, | il n'y a | aucun moyen | d'en saisir | les coupables, | incitant | une investigation interne | par | des examinateurs minutieux.
4. C L'administrateur de la société de vente en ligne détecte des données singulières dans le réseau interne.
- a. Un client | valide | les produit, | mais personne | dans le département informatique | ne peut | déchiffrer | comment, | même si | l'entreprise | qui vend le logiciel utilisé | a déjà été contactée.
 - b. Des clients | valides | les produisent, | mais personne | dans le département informatique | ne peut | déchiffrer | comment, | même si | l'entreprise | qui vend le logiciel utilisé | a déjà été contactée.
 - c. Un client | valide | les produit, | mais personne | dans le département informatique | ne peut | en déchiffrer | l'origine, | même si | l'entreprise | qui vend le logiciel utilisé | a déjà été contactée.
 - d. Des clients | valides | les produisent, | mais personne | dans le département informatique | ne peut | en déchiffrer | l'origine, | même si | l'entreprise | qui vend le logiciel utilisé | a déjà été contactée.
5. C Les élèves notent que la bataille d'Alésia est mentionnée plusieurs fois dans le livre scolaire.
- a. Le renvoi | explicite | les fait discuter, | mais le professeur | n'a pas le temps | de demander | avec quel résultat, | car il y a | beaucoup | de matières | à traiter | avant l'examen.
 - b. Les renvois | explicites | les font discuter, | mais le professeur | n'a pas le temps | de demander | avec quel résultat, | car il y a | beaucoup | de matières | à traiter | avant l'examen.
 - c. Le renvoi | explicite | les fait discuter, | mais le professeur | n'a pas le temps | d'en écouter | le résultat, | car il y a | beaucoup | de matières | à traiter | avant l'examen.
 - d. Les renvois | explicites | les font discuter, | mais le professeur | n'a pas le temps | d'en écouter | le résultat, | car il y a | beaucoup | de matières | à traiter | avant l'examen.
6. C Un couple brésilien lié au crime organisé est surveillé par le gouvernement argentin depuis longtemps avant de disparaître dans la jungle.
- a. Un agent | lâche | les recherche, | mais à présent | les sources | à l'intérieur du service secret | ne veulent pas | révéler | où, | de manière | à ne pas mettre | en péril | toute l'opération.
 - b. Des agents | lâches | les recherchent, | mais à présent | les sources | à l'intérieur du service secret | ne veulent pas | révéler | où, | de manière | à ne pas mettre | en péril | toute l'opération.
 - c. Un agent | lâche | les recherche, | mais à présent | les sources | à l'intérieur du service secret | ne veulent rien | en révéler, | de manière | à ne pas mettre | en péril | toute l'opération.
 - d. Des agents | lâches | les recherchent, | mais à présent | les sources | à l'intérieur du service secret | ne veulent rien | en révéler, | de manière | à ne pas mettre | en péril | toute l'opération.
7. C Trois étudiants emprisonnés sont graciés par le nouveau régime comme geste de réconciliation.
- a. La sentence | critique | les juge non-responsables, | mais les survivants | des victimes | de la terreur anti-gouvernementale | ne peuvent pas | comprendre | de quel droit, | protestant | contre | la décision | dans les rues | et sur Internet.

- b. Les sentences | critiques | les jugent non-responsables, | mais les survivants | des victimes | de la terreur anti-gouvernementale | ne peuvent pas | comprendre | de quel droit, | protestant | contre | la décision | dans les rues | et sur Internet.
- c. La sentence | critique | les juge non-responsables, | mais les survivants | des victimes | de la terreur anti-gouvernementale | ne peuvent pas | en comprendre | la justice, | protestant | contre | la décision | dans les rues | et sur Internet.
- d. Les sentences | critiques | les jugent non-responsables, | mais les survivants | des victimes | de la terreur anti-gouvernementale | ne peuvent pas | en comprendre | la justice, | protestant | contre | la décision | dans les rues | et sur Internet.
8. C Les sondages montrent à maintes reprises que l'extrême droite a pris le dessus, grâce à sa rhétorique contre la société multi-culturelle.
- a. L'opinion | aveugle | les avantage, | mais les commentateurs | plus prévoyants | demandent déjà | pour combien de temps, | d'autant | plus que | l'électorat | d'origine étrangère | augmente | d'année en année.
- b. Les opinions | aveugles | les avantagent, | mais les commentateurs | plus prévoyants | demandent déjà | pour combien de temps, | d'autant | plus que | l'électorat | d'origine étrangère | augmente | d'année en année.
- c. L'opinion | aveugle | les avantage, | mais les commentateurs | plus prévoyants | en contestent | la continuité, | d'autant | plus que | l'électorat | d'origine étrangère | augmente | d'année en année.
- d. Les opinions | aveugles | les avantagent, | mais les commentateurs | plus prévoyants | en contestent | la continuité, | d'autant | plus que | l'électorat | d'origine étrangère | augmente | d'année en année.
9. C Les jeunes sans-abris du campement ne veulent pas rester inertes plus longtemps.
- a. Un camion | détruit | les tente, | mais les anciens | de la petite communauté | sous-estiment | à quel point, | de sorte que | le conflit | avec les ferrailleurs | est déjà clairement | à prévoir.
- b. Des camions | détruits | les tentent, | mais les anciens | de la petite communauté | sous-estiment | à quel point, | de sorte que | le conflit | avec les ferrailleurs | est déjà clairement | à prévoir.
- c. Un camion | détruit | les tente, | mais les anciens | de la petite communauté | en sous-estiment | le danger, | de sorte que | le conflit | avec les ferrailleurs | est déjà clairement | à prévoir.
- d. Des camions | détruits | les tentent, | mais les anciens | de la petite communauté | en sous-estiment | le danger, | de sorte que | le conflit | avec les ferrailleurs | est déjà clairement | à prévoir.
10. C Les lionceaux couinent dans leur panique.
- a. Un volet | fixe | les piège, | mais les jeunes animaux | ne peuvent pas | comprendre | pourquoi, | jusqu'à | ce que | les braconniers | viennent | les ramasser | pour le transport.
- b. Des volets | fixes | les piègent, | mais les jeunes animaux | ne peuvent pas | comprendre | pourquoi, | jusqu'à | ce que | les braconniers | viennent | les ramasser | pour le transport.
- c. Un volet | fixe | les piège, | mais les jeunes animaux | ne peuvent pas | en comprendre | le but, | jusqu'à | ce que | les braconniers | viennent | les ramasser | pour le transport.
- d. Des volets | fixes | les piègent, | mais les jeunes animaux | ne peuvent pas | en comprendre | le but, | jusqu'à | ce que | les braconniers | viennent | les ramasser | pour le transport.
11. C Comme les invalides âgés n'ont pas d'argent pour vivre dignement, ils sont forcés d'endurer les défauts de la clinique laissée à l'abandon.
- a. L'infirmier | maudit | les mouche, | mais les vieux | se demandent | souvent | à quelle fin, | étant donné | que la plupart | des occupants | sont proches | de la mort.
- b. L'infirmière | maudite | les mouche, | mais les vieux | se demandent | souvent | à quelle fin, | étant donné | que la plupart | des occupants | sont proches | de la mort.
- c. L'infirmier | maudit | les mouche, | mais les vieux | s'en demandent | souvent | le sens, | étant donné | que la plupart | des occupants | sont proches | de la mort.
- d. L'infirmière | maudite | les mouche, | mais les vieux | s'en demandent | souvent | le sens, | étant donné | que la plupart | des occupants | sont proches | de la mort.
12. C Il y a du sang sur le sol du bureau, et les employés savent que c'est le sang du directeur détesté par tout le monde.
- a. La flaque | trouble | les calme, | mais le commissaire | et ses officiers | ne comprennent | pas vraiment | pour quelle raison, | même si | le tué | était décrit | comme un homme | agressif et menaçant.
- b. Les flaques | troubles | les calment, | mais le commissaire | et ses officiers | ne comprennent | pas vraiment | pour quelle raison, | même si | le tué | était décrit | comme un homme | agressif et menaçant.
- c. La flaque | trouble | les calme, | mais le commissaire | et ses officiers | n'en comprennent | pas vraiment | la raison, | même si | le tué | était décrit | comme un homme | agressif et menaçant.
- d. Les flaques | troubles | les calment, | mais le commissaire | et ses officiers | n'en comprennent | pas vraiment | la raison, | même si | le tué | était décrit | comme un homme | agressif et menaçant.
13. C Pour le stagiaire, les dossiers ne sont que des colonnes de chiffres cabalistiques.

- a. La technicienne | intègre | les signe normalement, | mais il est | toujours | complètement impossible | de prévoir | quand, | donc il ne | lui reste | plus qu'à | attendre silencieusement | sur sa chaise.
- b. Les techniciennes | intègres | les signent normalement, | mais il est | toujours | complètement impossible | de prévoir | quand, | donc il ne | lui reste | plus qu'à | attendre silencieusement | sur sa chaise.
- c. La technicienne | intègre | les signe normalement, | mais il est | toujours | complètement impossible | d'en prévoir | le délai, | donc il ne | lui reste | plus qu'à | attendre silencieusement | sur sa chaise.
- d. Les techniciennes | intègres | les signent normalement, | mais il est | toujours | complètement impossible | d'en prévoir | le délai, | donc il ne | lui reste | plus qu'à | attendre silencieusement | sur sa chaise.
14. C Les dames âgées sont très satisfaites de la jeune mariée et de ses demoiselles d'honneur.
- a. La robe | sublime | les charme, | mais comme d'habitude | elles invoquent | des raisons très différentes | pour expliquer | pourquoi, | chacune | impatiente | de mettre | son grain de sel.
- b. Les robes | sublimes | les charment, | mais comme d'habitude | elles invoquent | des raisons très différentes | pour expliquer | pourquoi, | chacune | impatiente | de mettre | son grain de sel.
- c. La robe | sublime | les charme, | mais comme d'habitude | elles invoquent | des raisons très différentes | pour expliquer | cette impression, | chacune | impatiente | de mettre | son grain de sel.
- d. Les robes | sublimes | les charment, | mais comme d'habitude | elles invoquent | des raisons très différentes | pour expliquer | cette impression, | chacune | impatiente | de mettre | son grain de sel.
15. C Le marché du quartier est connu pour les grandes côtes de bœuf qui sont livrées durant la nuit.
- a. Le boucher | sale | les tranche, | mais les clients | de la place Colbert | se demandent | incroyablement | quand, | vu que | la viande | est vendue | déjà marinée | et prête à l'emploi | en début de matinée.
- b. Les bouchers | sales | les tranchent, | mais les clients | de la place Colbert | se demandent | incroyablement | quand, | vu que | la viande | est vendue | déjà marinée | et prête à l'emploi | en début de matinée.
- c. Le boucher | sale | les tranche, | mais les clients | de la place Colbert | en demandent | incroyablement | la technique, | vu que | la viande | est vendue | déjà marinée | et prête à l'emploi | en début de matinée.
- d. Les bouchers | sales | les tranchent, | mais les clients | de la place Colbert | en demandent | incroyablement | la technique, | vu que | la viande | est vendue | déjà marinée | et prête à l'emploi | en début de matinée.
16. C Aujourd'hui, les enfants du vieux couple reçoivent un prix pour leur chanson dans une émission de télévision.
- a. Un modérateur | célèbre | les annonce, | mais les parents | ont des difficultés | à deviner | sur quel canal, | car le journal TV | semble | avoir disparu | une fois de plus.
- b. Des modérateurs | célèbres | les annoncent, | mais les parents | ont des difficultés | à deviner | sur quel canal, | car le journal TV | semble | avoir disparu | une fois de plus.
- c. Un modérateur | célèbre | les annonce, | mais les parents | ont des difficultés | à s'y | concentrer, | car le journal TV | semble | avoir disparu | une fois de plus.
- d. Des modérateurs | célèbres | les annoncent, | mais les parents | ont des difficultés | à s'y | concentrer, | car le journal TV | semble | avoir disparu | une fois de plus.
17. C Elodie pense que son mari a eu des aventures avec deux de ses assistantes.
- a. Le soupçon | grave | les espace, | mais l'insensible Gabriel | ne veut pas voir | encore | à quel point, | jusqu'à | ce qu'il | soit trop tard | pour sauver | le mariage.
- b. Les soupçons | graves | les espacent, | mais l'insensible Gabriel | ne veut pas voir | encore | à quel point, | jusqu'à | ce qu'il | soit trop tard | pour sauver | le mariage.
- c. Le soupçon | grave | les espace, | mais l'insensible Gabriel | ne veut pas en voir | encore | la sévérité, | jusqu'à | ce qu'il | soit trop tard | pour sauver | le mariage.
- d. Les soupçons | graves | les espacent, | mais l'insensible Gabriel | ne veut pas en voir | encore | la sévérité, | jusqu'à | ce qu'il | soit trop tard | pour sauver | le mariage.
18. C Pour la première fois, le chef d'orchestre ose faire jouer aux musiciens une pièce du compositeur hongrois farfelu.
- a. La notation | inverse | les rythme, | mais ni le directeur | ni les spécialistes invités | ne peuvent | expliquer | par quelle magie, | et les instrumentistes | se sont | complètement perdus | durant le concert.
- b. Les notations | inverses | les rythment, | mais ni le directeur | ni les spécialistes invités | ne peuvent | expliquer | par quelle magie, | et les instrumentistes | se sont | complètement perdus | durant le concert.
- c. La notation | inverse | les rythme, | mais ni le directeur | ni les spécialistes invités | ne peuvent | l'expliquer, | et les instrumentistes | se sont | complètement perdus | durant le concert.
- d. Les notations | inverses | les rythment, | mais ni le directeur | ni les spécialistes invités | ne peuvent | l'expliquer, | et les instrumentistes | se sont | complètement perdus | durant le concert.
19. C Les touristes sont encore traumatisés par la violence de la manifestation et veulent immédiatement retourner à l'hôtel.
- a. Une policière | sauve | les véhicule, | mais il est | un peu inquiétant | que personne ne lui ait indiqué | exactement | où, | vu qu'ils | sont logés | bien loin | en banlieue.

- b. Des policières | sauves | les véhiculent, | mais il est | un peu inquiétant | que personne ne leur ait indiqué | exactement | où, | vu qu'ils | sont logés | bien loin | en banlieue.
 - c. Une policière | sauve | les véhicule, | mais il est | un peu inquiétant | que personne ne lui ait indiqué | exactement | la longueur du trajet, | vu qu'ils | sont logés | bien loin | en banlieue.
 - d. Des policières | sauves | les véhiculent, | mais il est | un peu inquiétant | que personne ne leur ait indiqué | exactement | la longueur du trajet, | vu qu'ils | sont logés | bien loin | en banlieue.
20. C Les reliefs sophistiqués sur les colonnes du temple continuent de captiver les archéologues.
- a. Une enjolivure | précise | les orne, | mais les savants | n'arrivent toujours | pas à tomber | d'accord | depuis quand, | car le style | ne semble | correspondre | ni à une époque | ni à une autre.
 - b. Des enjolivures | précises | les ornent, | mais les savants | n'arrivent toujours | pas à tomber | d'accord | depuis quand, | car le style | ne semble | correspondre | ni à une époque | ni à une autre.
 - c. Une enjolivure | précise | les orne, | mais les savants | n'arrivent toujours | pas à tomber | d'accord | sur leur âge, | car le style | ne semble | correspondre | ni à une époque | ni à une autre.
 - d. Des enjolivures | précises | les ornent, | mais les savants | n'arrivent toujours | pas à tomber | d'accord | sur leur âge, | car le style | ne semble | correspondre | ni à une époque | ni à une autre.

SOI stimuli

1. Quels joueurs | du club de foot | a/ont vu | la femme | de mon frère | et où,/récemment, | si je puis me permettre?
2. Quels avocats | du cabinet | a/ont salué | le juge | du tribunal administratif | et quand,/fréquemment, | s'il vous plaît?
3. Quels vendeurs | du marché | a/ont dupé | le fils | de ma voisine | et quand,/ce jour-là, | si vous savez?
4. Quels acteurs | du théâtre | a/ont insulté | le critique | du quotidien | et comment,/l'autre jour, | si vous pouvez me dire?
5. Combien de maîtres | de kung-fu | a/ont vaincu | l'ours | du cirque | et comment,/dans l'histoire, | si je puis me permettre?
6. Combien de voitures | de course | a/ont dépassé | le gros | camion rouge | et quand,/en tout | s'il vous plaît?
7. Quels grimpeurs | de la région | a/ont chassé | l'ermite | du bois | et pourquoi,/le matin, | si vous savez?
8. Quels bandits | du désert | a/ont assiégé | l'armée | des Croisés | et quand,/en 1267, | si vous pouvez me dire?
9. Quels fidèles | de l'hindouisme | a/ont converti | le pasteur | de l'église paroissiale | et comment,/au final, | si je puis me permettre?
10. Quelles journalistes | de la chaîne | a/ont soudoyé | le président | du parti | et quand,/l'année dernière, | s'il vous plaît?
11. Quelle chanteuse | d'opéra | ont/a contacté | les musiciens | du groupe | et pourquoi,/en cachette, | si vous savez?
12. Quel chimiste | du groupe pharmaceutique | ont/a empoisonné | les collègues | du département | et comment,/par vengeance, | si vous pouvez me dire?
13. Quel escrimeur | de l'équipe nationale | ont/a salué | les adversaires | du concours | et quand,/avant la lutte, | si je puis me permettre?
14. Quel héros | de la gâchette | ont/a inhumé | les compagnons | de bordée | et où,/dans le désert, | s'il vous plaît?
15. Quel auteur | de romans policiers | ont/a consulté | les enquêteurs | du commissariat | et pourquoi,/par téléphone, | si vous savez?
16. Quel fonctionnaire | de l'Etat | ont/a protégé | les magistrats | du parquet | et quand,/en août dernier, | si vous pouvez me dire?
17. Quelle dompteuse | de lion | ont/a embauché | les cracheurs | de feu | et pourquoi,/illicitement, | si je puis me permettre?
18. Quel roi | de l'Angleterre | ont/a raillé | les ministres | de la Couronne | et pourquoi,/sans cesse, | s'il vous plaît?
19. Quel assassin | de la mafia | ont/a observé | les agents infiltrés | du service secret | et quand,/particulièrement, | si vous savez?
20. Quel explorateur | de l'institut ethnologique | ont/a caché | les autochtones | du village | et où,/quelque part, | si vous pouvez me dire?

Erklärung über Selbstständigkeit und Einhaltung wissenschaftlicher Standards

Hiermit erkläre ich, dass ich bei der Abfassung der vorliegenden Arbeit alle Regelungen guter wissenschaftlicher Standards eingehalten habe. Weiterhin erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe und über die Beiträge meiner Koautoren hinaus, welche in der beiliegenden *Erklärung über die Beiträge zu Gemeinschaftsveröffentlichungen* spezifiziert sind, keine Hilfe Dritter in Anspruch genommen habe.

Dario Paape

Berlin, November 2017