Tapping into semantic recovery : an event-related potential study on the processing of gapping and stripping

Ruijgrok, B.J.

Tapping into semantic recovery

An event-related potential study

on the processing of Gapping and Stripping

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: Trunks, by Esther Truijen

ISBN: 978-94-6093-289-2
NUR: 616

Tapping into semantic recovery

An event-related potential study on the processing of Gapping and Stripping

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 31 mei 2018
klokke 16:15 uur

door

Bobby Johan Ruijgrok

geboren 9 december 1972
te Leiden, Nederland

Promotores:	prof. dr. N.O. Schiller
	prof. dr. L.L.-S. Cheng
Co-promotor:	dr. C.L.J.M. Cremers
Promotiecommissie:	prof. dr. L. Frazier (Amherst University)
	prof. dr. J. Grijzenhout (secretaris)
	dr. L. Pablos Robles
	prof. dr. F.N.K. Wijnen (Utrecht University)

The beautiful lull
The dangerous tug
We get to feel small
From high up above
And after a glimpse
Over the top
The rest of the world
Becomes a gift shop

[Downie, Fay, Langlois, Baker, and Sinclair (1996)]

Contents

Acknowledgements

Doing a PhD isn't particularly easy and my mum's prosodic phrasing of *Waar begin je aan?* was spot on. But since I've been surrounded and supported, either knowingly or accidentally, by a bunch of nice people, I completed this dissertation in due time – give or take.

This work wouldn't be here without Lisa Cheng. She successfully chased me up in 2011 to write and submit a proposal to the Netherlands Organisation for Scientific Research (NWO). On the basis of an earlier draft by Crit Cremers and Niels Schiller, I gladly did so. I thank all three of you, being a well-balanced team of *promotores*, for your confidence in me and the great amount of freedom that you have given me.

Prior to submitting a proposal, I was introduced to the technique of ERPs by Leticia Pablos Robles during a joyful internship. *Domo arigato gozaimasu!* Following that, Jirina Vandenbussche accompanied me in the course of writing, which was helpful to keep a steady tempo. *De Beebie!* Then, I approached my first ever job interview for which I was properly prepared after a serious mock interview with Maghiel van Crevel. *Dra!*

From practically day one, while picking up keys to the office, I became friends with Olga Kepinska and, soon after, her office mate Eleanor Dutton. The three of us would forge a mutual supervisorship helping each other to escape the valleys of (research) misery and to significantly improve our experiments and drafts. Olgi, you've been an outstanding travel companion to conferences and summer/winter schools. Elly, I value greatly the combination of lab managing, designing and lecturing a great methodological course, and musical intermezzos. Also, thanks for proofreading these 160 pages. Thank you witty and creative guys for making this PhD trip a blast.

Shortly after I started my project I learned that, especially when dealing with practicalities regarding labs and office, essential to a work space is support staff – the *WD-40* of an organisation. Thanks to regular advice, convos and banter from and with Jaap van Belen, Brahim Boulahri, Betty de Jonge, Adri van de Kerkhof, Alex van der Meer, Jos Pacilly, Mo 'Pakkie' Rbii and

Merel van Wijk, I found numerous shortcuts to overcome unforeseen hiccups, which helped me to stay sane in the workplace.

I carried out my first two ERP experiments in the basement of the Psychology Department. I am thankful to Guido Band, Maureen Meekel and Kerwin Olfers for their tips and tricks.

In the meantime at LUCL, I got to know many staff members, PhD candidates, post-docs, students and guests, who all contributed to a warm and stimulating work environment. Special memories are connected to my first office that was next to a natural common room. There, I shared many laughs and chats with Linda Badan, Jermy Balukh, Martine Bruil, Anne van der Kant, Carmen Ebner, Andreea Geambasu, Allison Kirk, Martin Kohlberger, Marie-Catherine 'MC-ke' Michaux, Borana Lushaj, Marieke Meelen, Laura Migliori, Jessie Nixon, Victoria Nyst, Gareth 'Gary' O'Neill, Dafina Ratiu, Benjamin 'Bennie' Suchard, Mulugeta Tsegaye, Marijn van 't Veer, Daan van de Velde and the incidental LUCL passerby.

In the second leg of my PhD, I carried out three ERP experiments and finalised my thesis during which I received considerable moral support from Kate Bellamy, Saskia 'Saksia' Lensink, Viktorija 'Happiness Agent' Kostadinova (*coolage!*), Morana Lukač, Min Liu, Nazar Nazarudin, Matthias Pache, Alex Reuneker, Gulnaz 'Guli' Sibgatullina, Yang Yang, Cesko 'Ceski' Voeten and Man 'Captain Slow' Wang. Furthermore, I attended the LSA in Chicago during which I broadened my views through Klinton Bicknell, Lyn Frazier, Philip Hofmeister, Ellen Lau, Roger Levy, Colin Phillips and Ming Xiang. Conversations during drinks, meals, walks and bicycle rides I had with Aaron Apple, Iria de Dios-Flores, Lara Ehrenhofer, Rabia Ergin, Ophelie Gandon, Maren Berg Grimstad, Heidi 'Birthday Buddy' Getz, Kelly Kasper-Cushman, Jo Mackby, Paula Márquez Caamaño, Zoey Liu, Joey Lovestrand, Darcy Rose and Jolien Scholten still resonate.

I welcomed a total of 186 participants during the past five years. Wouter Broos, Isabella Jordanoska, Judith Kelholt and Naomi Nota played a crucial role in reaching this amount.

Although the last five years I have neglected a large group of friends, a small group still kept an eye on me: Petra & Jeroen van Alphen, Dirk & Gebke van den Bosch, Vincent & Trudy van Diemen, Niels Joor, Hans Lardee, Nicoline & Igor Sekeris, Jacco Spaargaren, Gijs van Straalen and Koen de Vries. I promise there will be make-up time for everyone I missed.

I wouldn't be where I am now without the unconditional love, humour and nurture of my parents Han & Elly. During my PhD, the family expanded by 1/3 as I happily became uncle of two jubilant little fellows – thanks to Judith & Yorick.

Finally, but never least, I find myself *kei*-lucky waking up everyday next to my doting 'best' half. Whether it's rebuilding an entire house, starting a business, completing a PhD, or changing the walk of life otherwise, Apie, your bright and high-spirited moves make anything go lightly. I'm looking forward to our next adventure.

Abbreviations

AFF	affirmative
A-P	articulatory-perceptual
CCG	Categorial Grammar
C-FOCUS	contrastive focus
C-I	conceptual-intentional
CS	conceptual structures
DET	determiner
F	feminine
HPSG	Head-driven Phrase Structure Grammar
INF	infinitive
IntP	intonational phrase
LF	logical form
M	masculine
NEG	negation
NP	noun phrase
ORPH	orphan
PF	phonological form
PP	prepositional phrase
PREP	preposition
PS	phonological structures
PST	past
QUD	question under discussion
REFL	reflexive pronoun
SG	singular
SS	syntactic structures
VP	verb phrase

CHAPTER 1

Introduction

1.1 The nature of this study

This dissertation is an example of interdisciplinary experimental linguistic research. The pivotal aim is to connect theoretical linguistic insights with behavioural and neuroscientific data. While PhD research in the natural sciences is normally concluded with a collection of (submitted) peer-reviewed journal publications, the motivating force behind this dissertation has been to compose a long essay – a book. In that sense, the title of the grant which has fuelled this research, *Promoties in de Geesteswetenschappen* (PhDs in the Humanities) provided by the Netherlands Organisation for Scientific Research, has been entertained quite formally. However, to a certain extent this thesis differs from a typical linguistic dissertation. Some chapters may resemble the form of a journal article reporting experiments where footnotes are scarce.

1.2 Interpretation of elided structures: some basic concepts

In spoken and written language, there are often cases where words that can be understood from contextual clues can be omitted. For example, we are able to interpret the second clause *Jerry a bike* in (1) as meaning that Jerry stole a bike, even though the verb *stole* is not physically present in the second clause.

(1) Tom stole a car, and Jerry a bike.

Jerry a bike is linguistically speaking not a proper clause, but an incomplete linguistic structure. Yet, we are able to understand that *Jerry* **stole** *a bike*; he did not, for example, buy one. We term this phenomenon "ellipsis". Ellipsis is arguably the most prominent example in human language of compromised mapping between linguistic form and meaning. When we study ellipsis, we aim to understand how it is possible for language users to arrive at an interpretation in the absence of form. While there are several ellipsis types (I refer the interested reader for a concise introduction to Merchant, 2017)[1], this dissertation employs the ellipsis type "Gapping"and its sub-type "Stripping", of which we see examples in (1) (Gapping) and (2) (Gapping and Stripping). Elided elements are denoted by <*e*>.

(2) a. Eva bought a book, and Agnes <*e*> a CD. (*Gapping*)
b. Eva bought a book in the shop, and Agnes <*e*> in the supermarket. (*Gapping*)
c. Eva bought a book in the shop, and Agnes <*e*> too. (*Stripping*)

As can be observed, Gapping-like constructions are characterised by an omission of at least the finite verb in the second conjunct of a coordinate structure. The remaining phrases in the second conjunct – called "remnants"– contrast with their correlates in the first conjunct. Gapping involves at least two remnant phrases, Stripping involves one remnant and an additive marker ("too"). Crucially, we are able to recover the meaning of the omitted material – called the "antecedent" – in order to fully interpret the right conjuncts in (2). We use information that we retrieve from the left conjunct and we integrate this information in the right conjunct (also sometimes referred to as "reduced" conjunct).

Ellipsis might be conceived of as an "anaphoric" relation between an antecedent and omitted structure. However, in contrast to overt anaphoric relations such as those constituted by pronouns (he, she, etc.) and reflexives (himself, herself, etc.), ellipsis lacks overt form. For example, the reflexive *herself* in (3) is overt linguistic material that refers to the antecedent *Sheila*.

(3) Sheila saw herself in the mirror.

It is important to realise that in order to understand the elliptical constructions in (1) and (2) the interpreter can only use the antecedent within the linguistic context. Gapping (and Stripping), by definition, "requires a verbal context" (Cremers, 1993:117); that is, linguistic material is required in the process of interpretation. Listeners (or readers) somehow need to retrieve the intended proposition, and the missing information is provided by the left conjunct. Therefore, the ellipsis type under investigation in the current study is not just an instance of so-called "underspecification", which is abundant in human

[1] I follow Merchant (2017) in categorising Gapping as an ellipsis type (see for contrasting ideas Johnson, 2009; Lappin & Benmamoun, 1999)

language. For example, to interpret a sentence such as (4), we add information which cannot be inferred on the basis of the sentence alone.

(4) She waited there but he didn't show up.

Without an antecedent within the *linguistic* context we make use of *extra-linguistic* context, which helps us to capture, for instance, time and place. In this thesis I will be concerned with elliptical structures that can only be resolved within a linguistic context.

The theoretical literature on ellipsis is generally concerned with questions regarding the conditions under which ellipsis is permitted (or "licensed") and the level of description at which the relation between antecedent and ellipsis site should be formalised. For example, some scholars emphasise the import of syntactic operations, while others favour a semantic perspective. Prosody is also considered to be an important factor. In Chapter 2, I will examine these issues with reference to Gapping-like constructions. It will appear, that to account for distributional properties of Gapping and Stripping, a successful account should combine syntactic, semantic and prosodic factors. By extension, recovery strategies that are employed to resolve ellipsis include these factors. The question remains: what is the division of labour between syntactic, semantic and prosodic-based mechanisms?

While Gapping (and other ellipsis types) have been studied extensively in the theoretical literature, the present study investigates the neurophysiological processes that are at work to resolve "gapped" or "stripped" elements such as *bought* in (1) and *bought a book* in (2b) and *bought a book in the shop* (2c).

1.3 Levels of analysis: grammar, processing and neurons

How do we connect theoretical concepts to processes that take place in the brain? While neurophysiological literature on ellipsis is still in its infancy, many theoretically-oriented scholars have explored how theoretical constructs might be realised cognitively. At least within the Generative enterprise, formal theories of linguistic structure are theories of *competence*: the abstract mental speaker-hearer's knowledge of a language – the finite set of rules for producing and comprehending an unlimited amount of utterances. Chomsky (1965) was the first to contrast this with the notion of *performance*: the actual utterances produced by speakers. Theoreticians, including those who work on ellipsis, try to formulate conditions that explain the "grammaticality" of a certain construction – without being interested *per se* in how these conditions are accessed during language production and comprehension. This has consequences for the way theories have been developed, but also for experimental researchers, who need to be able to refer to theory, but whose methodology concerns the actual production and comprehension of language. One of the objectives of this dissertation is to try to bring together the best parts of both the theoretical and experimental worlds.

Both theoreticians and experimental researchers conceive of the speaker-hearer's mental knowledge of language in terms of mental representations. I will use the definition as proposed by Marr (1982):

> A representation is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this.
>
> [Marr (1982:69)]

Like cognitive scientists, I will use the word "representation" as referring to a psychological object. I assume that processes of human behaviour, cognition, are guided by computational procedures which operate on and amount to internal representations. Computation requires a "grammar", i.e. rules for a combinatorial mechanism, in order to apply unification. It is widely accepted that the human language system exploits a "mental lexicon", a mental dictionary containing information regarding a word's syntactic characteristics, meaning and pronunciation. There is, however, less agreement as to the level of detail contained in the lexicon. Despite this we may assume a computational procedure during which a word is retrieved from the mental lexicon to be integrated with linguistic material processed earlier. For example, in sentence (2a), the items *buy* and *a book*, both having representations at a lexical level, can be unified to form a representation of a verb phrase. How does such a representation look in the right conjunct *Agnes a CD*? In ellipsis research, as we will see, much has been written about representation and computation. Broadly speaking, syntax-oriented accounts hold that the interpretation of elliptical structures depends on the reconstruction of syntactic structure or a copy thereof. A representation of a fully-fledged syntactic structure would resemble a "surface" structure, which would be pronounced if it was not elided. The ellipsis site reflects the syntactic identity of the antecedent. This contrasts with semantics-oriented accounts that emphasise the role of (rules at) a conceptual level of representation. Here, the idea is that ellipsis is resolved at a "deeper" level of representation, rather than by means of a representation of unpronounced surface structure. Syntactic accounts emphasise the notion of parallel syntactic structure, while semantic-oriented accounts are concerned with parallel properties of more fully interpreted chunks.

Translating theory into procedural (or "processing") terms may not always be straightforward, however. With reference to the implementation of such processes in the brain (i.e. how the representation and computation of elliptical structures is executed at a neuronal level), little has been achieved so far. The current study hopes to contribute to all three levels – grammar, processing and implementation – because it is my conviction that language should be understood at all these levels.

1.3.1 Grammatical levels of analysis

This research takes as starting point that a sentence - either written or spoken - is a pairing of form and meaning. The form of a sentence comprises the actual output (orthographic when written, phonetic when spoken) and its structural properties. A framework that has proponents in both linguistics and psychology is the "tripartite parallel architecture of the grammar"as depicted in Figure 1.1 (see Jackendoff, 1997 and subsequent work). In this framework, three distinct levels of representation - phonological (output), syntactic (structural), and semantic (meaning) - are assumed to be components that are governed by principles and rules of their own. Interface modules are included to specify the links between the parallel components.

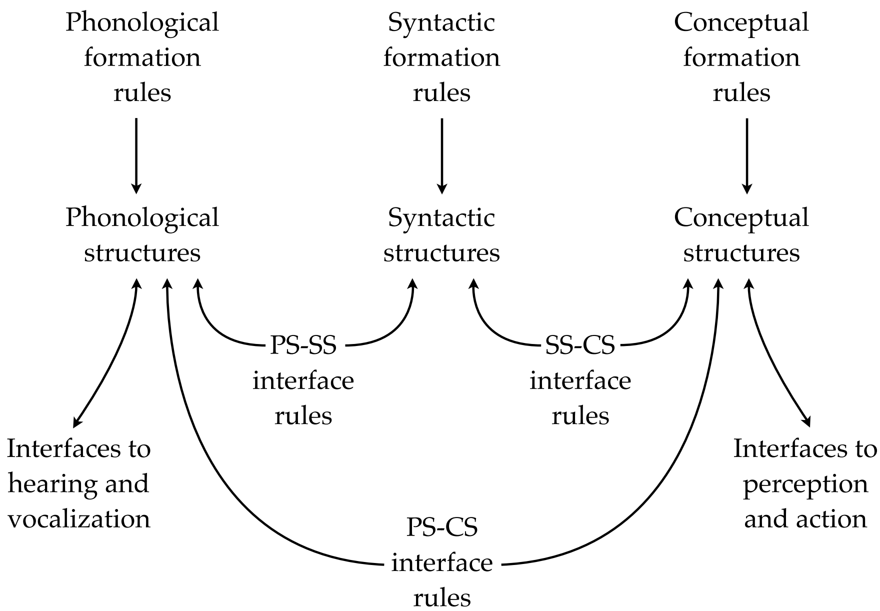


Figure 1.1: The parallel grammar architecture (Jackendoff, 2002:125). Phonological structures is abbreviated as PS, syntactic structures as SS, and conceptual structures as CS.

In the spirit of parallel architecture I acknowledge the independent combinatorial character of three information types: phonology, syntax, and semantics. The rationale behind this view is the assumption that, in contrast to, for example, Minimalist approaches (initiated by Chomsky, 1993), phonological and semantic representations are not exclusively derived from syntactic structure but rather constrained by it. In that sense the Chomskyan tradition is unidirectional and syntax-centred, as can be seen in Figure 1.2 which represents a modern version of a minimalist model. Lexical items are combined through syntactic formation rules. At some point during the derivation, the computation is split and interpreted at the components Phonological Form (PF) and Logical Form (LF); this point is known as "Spell-Out". PF interfaces with an Articulatory Perceptual (also referred to as "sensorimotor") system while LF interfaces with a Conceptual Intentional system. Since this is a competence model, it is only concerned with the syntactic derivation up until the components PF and LF, culminating in representations at these interpretative levels. Note that LF cannot be seen as a semantic level as it is an intermediate representation for matching syntactic structure with rules of interpretation. Although Spell-Out is assumed to occur throughout the derivation in cyclic "phases", a derivation does not represent processing steps. Since this model does not make claims about the way PF and LF communicate with the output systems, it is not straightforwardly utilised outside the field of syntactic analysis. Nonetheless, Minimalist approaches have contributed a great deal of work on ellipsis.

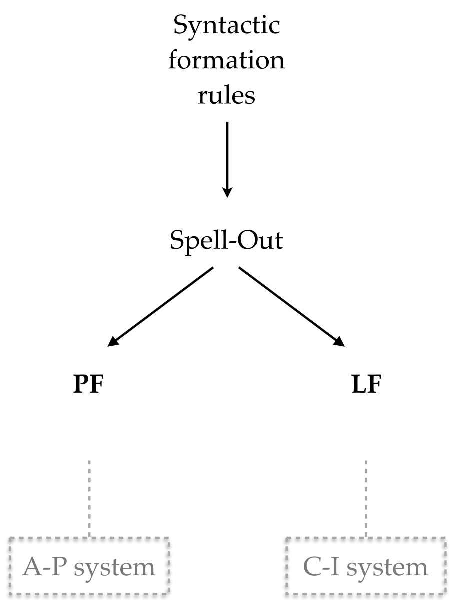


Figure 1.2: A minimalist model of the grammar (after Hornstein et al., 2005:73). Phonological Form is abbreviated as PF, Logical Form as LF, Articulatory Perceptual as A-P and Conceptional Intentional as C-I. The model is silent as to how PF and LF interact with the output systems.

Whereas in the Minimalist approach syntactic operations are core and inherent, in Jackendoff's model phonological and semantic structures are subject to combinatorial rules of their own levels, which may produce structures that, apparently, do not have a one-to-one correspondence with syntactic structures (SS). In (5) we see an example of a mismatch between intonational and syntactic structures. The bracketing in (5b) and (5c) represents possible intonational phrasings (denoted by IntPs) that do not always converge with syntactic constituents as represented in (5a), while no semantic difference between (5b) and (5c) can be identified.

(5) a. [$_{NP}$ *Sesame Street*] [$_{VP}$ is [$_{NP}$ a production [$_{PP}$ of [$_{NP}$ the Children's Television Workshop]]]]
b. [$_{IntP}$ *Sesame Street* is a production of] [$_{IntP}$ the Children's Television Workshop]
c. [$_{IntP}$ *Sesame Street*] [$_{IntP}$ is a production] [$_{IntP}$ of the Children's Television Workshop]

[Jackendoff (2002:118-119)]

Jackendoff notes that intonational contours do not follow syntactic phrases at all times. In other words, phonological rules seem to apply to phonological constituents that are not always exact mappings of syntactic categories; the intonational bracketing is governed by independent phonological rules which may apply independently of syntactic rules. While minimalist approaches have tried to circumvent intonational-structural mismatches (see for example Dobashi, 2009), Jackendoff (1997, 2002) reasons that conceptual structures may be governed by independent rules too, as he demonstrates with the famous example (6) from Chomsky (1957). Although neither of the sentences make sense, native speakers of English have the intuition that (6a) is grammatical and (6b) is not.

(6) a. Colorless green ideas sleep furiously.
b. Furiously sleep ideas green colorless.

[Chomsky (1957:15)]

While structurally (6a) is correct, the mechanism that determines the nonsensical status of this sentence must be sought at the level of conceptual structures. In addition to the above example, Jackendoff (1997:33-35) puts forward other phenomena to show that the mapping between syntactic structure and conceptual structure is not a one-to-one relation. It should be noted that Conceptual Structure (CS) is not assumed to be part of the language faculty per se, though the SS-CS interface rules are part of the language system. Crucially different from a Minimalist approach, this framework specifies links to enable the language system to interact with processes of logical and heuristic reasoning. Jackendoff considers the language faculty as consisting of levels of representation as shown in Figure 1.2, yet, they are assumed to interact and to have combinatorial power of their own. The level of syntax is assumed to consist of "syntactic formation rules" that play a mediating role – rather than a deterministic role – between linearly ordered phonological strings and the linearly unordered structure of meanings.

The advantage of the tripartite framework is twofold. Firstly, in the following chapters we will see that ellipsis is a multidimensional phenomenon in which syntactic, semantic and phonological constraints apply. Theoretical approaches usually take one of these dimensions as a starting point. The model can be used to evaluate and compare a variety of approaches, which may be embedded in different frameworks. Secondly, it may be used as an umbrella instrument to evaluate theoretical and experimental hypotheses, taking it as a starting point for an integrated framework accommodating theory and neurocognitive data. In particular, the architecture can be extended as a processing framework that integrates a crucial role for working memory (Jackendoff, 2002:196-200). Although it is not always explicitly integrated in theories of sentence processing, I assume that working memory plays an important role in language use.

Formation rules of the three components (phonology, syntax and semantics, in other words, the "grammar"), as depicted in Figure 1.1, may be attached to corresponding processors. Jackendoff refers to this linkage as an integrative process: "For each set of formation rules that defines a level of linguistic structure, the language processor requires an integrative process that uses these principles to construct structures at that level" (Jackendoff, 2002:198). Likewise, interface constraints guide corresponding processors that link the separate levels, as sentence comprehension (and production for that matter) consists in the integration of all levels of analysis. During language use, the integrative process is sustained by a linguistic working memory component that is to be understood as a "dynamic workbench" where three independent processors work in parallel, assembling and integrating linguistic structures of different levels (Jackendoff, 2002:200).

1.3.2 Sentence processing

While the tripartite architecture encompasses both language production and comprehension, this thesis focuses on the latter. There is general agreement that during listening and reading representations are built incrementally and that the sentence comprehension processor has four main tasks:

- retrieve grammatical and lexical information of incoming words
- analyse the grammatical structure identifying each word's position in the sentence; this is known as syntactic parsing
- analyse the prosodic structure identifying clause boundaries and the relative contrasts between phrases; this is known as prosodic parsing
- combine individual words, phrases and prosodic information to yield a representation of the meaning of a sentence; this part is referred to as semantic interpretation.

These processes of retrieval and integration are immediate, automatic and "it appears that there is no measurable lag between recognising a word and attempting to integrate it into a sentence-level syntactic and semantic representation" (Staub, 2015:204). On a word-by-word basis, the processor parses each new incoming word to retrieve the necessary information. Incrementally, the processor postulates phonological, syntactic and semantic representations to construct the meaning of a sentence. There is, however, less agreement as to the autonomous status of different information types and the way they interact. For example, to what extent are a word's syntactic features processed separately from its semantic information?

Historically, parsing is related to psycholinguistic "syntax-first models" of sentence processing: a syntactic structure is constructed serially using word-category information, independently of lexical-semantic information, which is processed at a later stage. The combination of the individual words and phrases results in a representation of the sentence's meaning. As a consequence, semantic interpretation relies on structure building. If an initial syntactic structure cannot be completed, reanalysis may take place. Frazier's research, starting in 1978 in collaboration with Fodor but revised in later (still ongoing) work, is grounded on this conception. In their proposed "Sausage Machine model", a syntactic analysis is constructed by means of funelling incoming information through a window of roughly six words, at which point the parse is clipped off (visualise the sausage machine here) and passed to a second stage to complete the interpretation. To a large extent Frazier adheres to a modular position as proposed by Fodor (1983). Influenced by Chomsky's ideas, Fodor apprehends language as an encapsulated input system operating on domain-specific information structures. The output of this system is then open to further evaluation by a general cognitive system. The fact that every module is assumed to be impenetrable means that no other cognitive process may affect its operation. Nevertheless, Frazier's recent position appears to be to arguing for a rather dynamic approach in which a competence module and performance module may interact (Frazier, 2015).

Contrasting serial models, some scholars have proposed that different types of information are processed incrementally in parallel in an interactive way. While a less differentiated representational vocabulary is assumed, interplay among different types of information is far less constrained. For example, lexical or more global semantic information is assumed to be able to influence structure building right away. Tyler and Marslen-Wilson (1977) could be taken as the starting point of this line of research, which has evolved to give us interactive "constraint-based models" in which (for example) syntactic constraints may be overruled by semantic constraints. Note that, just as the syntax-first tradition, this line of research acknowledges the notion of hierarchical structure. The outer end of parallel models, however, argues that the relation of words in a sentence can be explained in statistical terms – word order based on probabilities. Within the framework of so-called "connectionism" it has been proposed that language use is characterised by domain-general "low-level" processing units (at the neuronal level), in disagreement with the notion of "high-level" (abstract) symbolic representations proposed by formal linguists (see for example Rumelhart & McClelland, 1986).

The concept of probability is quite common in the psycholinguistic literature, where frequency effects on language processing are commonly reported. For example, the identification of a given word is much faster if it is a frequently used word (see for a discussion Bradley & Forster, 1987). In that sense, word frequency can help to build a probabilistic word model. As such, frequency effects are expected to be important during sentence comprehension. And not only at a lexical level: on the basis of syntactic frequency, the parser may predict certain structures. Some scholars have modelled this phenomenon in terms of probabilities (see for example Hale, 2011; Levy, 2008). However, to date there is, to my knowledge, no processing model that can account for syntactic, semantic, and prosodic phenomena including predictions. As has been suggested by Gibson and Pearlmutter (1998), such a model would be both parallel and constraint-based in nature.

Jackendoff's architecture may be seen as compromise between nativists (Chomskyan tradition) and behaviourists (connectionist approach). On the basis of "structure-constrained modularity" he adopts a flexible version of Fodor's proposal (Jackendoff, 1997:219). Like Fodor, he regards the brain "as a collection of specialists rather than an all-purpose cognizer", yet modules may interact through interface rules (as can be seen in Figure 1.1). While the distinction between the three levels of representation in Jackendoff's model may be module-like, it does not imply that each module has a one-to-one encapsulated mapping in the brain, which is rather dynamic. For example, Poeppel and Embick (2005) have argued that parts of cortical networks at work during syntactic computations may also be involved in phonological processes. While both nativists and behaviourists would agree that a representation of a linguistic unit, like *Agnes a CD* in (2a), is a mental state that is reflected by the activation of some group of neurons in the brain, the debate surrounds the degree to which it is guided by general principles; in other words, to what extent is a representation guided by language-specific rules?

1.3.3 Event-related brain potentials

Electroencephalography (EEG) is a method used to record electrical activity in the brain. Neurons communicate by pumping ions from one to another. In large amounts, ionic current flows result in tiny changes in electric potential when compared with a reference point. This amounts to a tiny change in voltage, which can be measured in microvolts. EEG enables us to measure the exact point in time when such changes in ion exchange occur.

Typically, EEG data can be analysed as event related potentials (ERPs). This means that the EEG signal ("potential") is inspected relative to specific time points ("events") in the experimental presentation. For example, when experimenters mark the time point at which a stimulus appears on-screen (this is referred to as "time-locking"), they can then analyse how the brain activity responds to that particular stimulus. The EEG responses are averaged across stimuli and participants; these average responses that are time-locked to such a stimulus are what is known as ERPs. It is a common methodological approach for tackling questions regarding the nature of semantic, syntactic and (to a lesser extent) prosodic processes and how they interplay.

In response to distinct experimental manipulations, discrete ERP patterns (also known as "ERP components") have been found. For example, experimenters may manipulate a syntactic characteristic of sentences to see to what extent this manipulation causes differences in the EEG waveforms in terms of polarity (positive of negative), latency (onset and duration of a deflection) and distribution (topographic reference). ERPs that have been consistently identified in the literature are typically given names according to the polarity and onset. Typically, "P" and "N" refer to positive or negative (note that this does not imply that positivity is 'good' and negativity is 'bad'!). Numbers of a component refer to the time point in milliseconds when the potential is observed after stimulus onset. For example, a negative component around 200 ms after stimulus onset is called an N200 or N2. The duration and distribution further help to determine the relationship of the ERP to underlying cognitive processes.

Five main markers have been identified in the literature with respect to language processing: Closure Positive Shift (CPS), Early Left Anterior Negativity (ELAN), Left Anterior Negativity (LAN), Negative 400 (N400), and Positive 600 (P600). Table 1.1 below lists the five ERP components categorising their latencies, distributions, and relationships to linguistic processes.

Friederici (2002) is an example of a sentence processing model that is based on findings from ERP data. This model of auditory sentence processing aligns very much with a syntax-first approach while linking the sequential processing steps to distinct brain sites connected to working memory. As can be observed in Table 1.1, the ERP findings can be formulated as a serial procedure starting with (superficial) syntactic structure-building on the basis of a word's category, after which interpretation may take place. At different times during this process prosodic information may be deployed. However influential in

Component	Latency (ms)	Distribution	Linguistic process(es)
CPS	0-600	bilateral centro-parietal	phonological/prosodic phrasing
ELAN	120-220	either bilateral or left anterior	syntactic structure building and phrase structure violations
LAN	300-500	either bilateral or left anterior	processing of semantic relations and morphosyntactic violations
N400	around 400	centro-parietal bilateral often with a slight right hemisphere focus	processing of conceptual/semantic information
P600	300-900	centro-parietal (fronto-central related with complexity)	wide variety of syntactic violations, syntactic reanalysis and repair, retrieval, increased syntactic complexity and ambiguity, syntactic and semantic integration

Table 1.1: Main ERP components related to linguistic stimuli: CPS (Closure Positive Shift), ELAN (Early Left Anterior Negativity), LAN (Left Anterior Negativity), N400 (Negative 400), P600 (Positive 600) and their latencies, distribution and relation to linguistic processes (after Friederici et al., 2002; Gouvea et al., 2010; Steinhauer, 2003; Swaab et al., 2012).

the field of neurolinguistics, the strictly serial nature of Frederici's model has been criticised by subsequent proposals which promote parallel or interactive procedures (see for example Hagoort, 2005; Hickok & Poeppel, 2004). In a later version, Friederici (2011) does endorse a comprehension process that consists of "several subprocesses that take place in a serial cascading and partly parallel fashion" encompassing neuronal pathways supporting sound-to-motor mappings and higher-level language processes.

Not everybody agrees that all components listed in Table 1.1 reflect processes specific to language. For example, it has been argued that the P600 belongs to the P300 family known to reflect domain-general phenomena such as context updating and surprise effects of unexpected stimuli (see for an initial discussion Coulson, King, & Kutas, 1998; Osterhout & Hagoort, 1999). Gouvea et al. (2010) suggest that a P600 may reflect retrieval and relation-forming processes but that it depends on the onset and duration. While their study was not intended to take a stance in the "P600 as P300" debate, they note that their account "could be extended to a domain-general account of the P600" (Gouvea et al., 2010:183).

More recently, the P600 generated by frontally-oriented neuronal activity is assumed to be a reflection of integration processes proper, that is, the relative difficulty in establishing a coherent utterance representation (Brouwer, Crocker, Venhuizen, & Hoeks, 2016; Brouwer & Hoeks, 2013). In this account, the N400 amplitude is assumed to exclusively reflect the relative difficulty of retrieval of lexical information from memory, contrasting with others who additionally relate semantic composition or integration to the N400 effect. Coined as the "Retrieval-Integration" account, it is underpinned by a neurocomputational model that successfully simulates ERP modulations in semantic processing. While an analogous process is required for ellipsis resolution, it is an open question to what extent this account can be extended to ellipsis data, given that an antecedent for ellipsis is retrieved from an earlier interpreted chunk rather than from lexical memory.

Since the details of ellipsis processing models and their relation to biologically plausible neurocognitive models of language comprehension will be examined later on, the purpose of this section is to lay out a road map of sentence processing in relation to grammar and ERPs. Noting different starting points, syntax-oriented and semantics-oriented, I would like to make clear that I do not take a position *a priori*, for the reason that research groups having a clear *a priori* preference for model *X* tend to provide evidence in favour of model *X* more often than not. For example, data may be analysed in such a way that a statistically significant result is forced to occur. This phenomenon is also known as "confirmation bias". Although the existence of confirmation bias in science has been acknowledged and suggestions have been made to prevent it, it remains a delicate issue (see for example MacCoun & Perlmutter, 2015; Nickerson, 1998). I would like to avoid any such 'predisposition', though I will follow the generally accepted notion of that phonological, syntactic and semantic structures are built incrementally during listening and reading.

1.4 Outline of this dissertation

This dissertation consists of two main parts. After the general introduction provided in this chapter, I discuss in Chapters 2-4 the relevant theoretical and experimental background on Gapping and Stripping in which it is shown that they have a multidimensional character. This provides us with a well-grounded starting point for the experiments that are reported in Chapters 5-8.

The theoretical accounts that I review in Chapter 2 can be broadly categorised as syntax-oriented and semantics-oriented. Syntax-oriented accounts emphasise the requirement of structural parallelism between antecedent and ellipsis and generally hold that the interpretation of elliptical structures depend on the reconstruction of syntactic structure or a copy thereof. This contrasts with semantics-oriented accounts that propose that interpretation is done by referral to (rules of) a conceptual level of representation. Although this rather simplistic differentiation between syntactic and semantic accounts has been the driving force behind the current project – linking this differentiation to electrophysiological data – I argue that Gapping-like constructions cannot be captured in either syntactic or semantic terms. An additional level of analysis, prosody, is discussed.

Chapter 3 covers experimental literature on ellipsis which reflects characteristic issues raised by the theoretical literature to a certain extent. Two behaviourally motivated parsing models that are grounded in theoretical insights are taken into consideration and are proposed as a possible link between theory and data. The proposal "Copy α" is inclined toward syntactic-oriented accounts. This contrasts with a "cue-based mechanism" that leans towards semantic-oriented accounts. Again, the role of prosody is examined, as well as the relevant ERP components that have been found in relation to the recovery of elliptical structures.

In Chapter 4, I argue that a mapping between existing theoretical insights and actual processing may not always be straightforward or even justifiable. Nonetheless, I arrive at a comparison of Copy α and the cue-based mechanism with respect to the timing of processes of retrieval and integration. By doing so, I can utilise these mechanisms to make hypotheses for the subsequent ERP experiments. Since individual differences may lead to differences in (amplitudes of) ERP components and may be ascribed to natural variability in the capacity of human working memory, I propose a suitable working memory test.

Chapter 5 starts with a report of a replication study on verb Gapping in Dutch. On the basis of stimuli used in this replication study, I designed and pretested new Gapping and Stripping stimuli for this dissertation. The method and results of the pretests are also reported in Chapter 5. In each of the following chapters, I test a representational dimension separately: syntax in Gapping and Stripping in Chapter 6, semantics in Stripping in chapter 7 and prosody in Gapping in chapter 8. Overall, I aim to estimate the relative import of these dimensions during the resolution process of Gapping-like constructions.

In Chapter 6, I report two ERP experiments on Gapping and Stripping constructions in which I modulate structure in the right conjunct and in the left conjunct. I hypothesise that modulation of structure would be reflected by early ERPs related to retrieval of a fully-fledged syntactic structure as a reflection of a Copy α mechanism. As an alternative, I suggest that a cue-based account predicts relative ease of retrieval but a relatively more costly integration process. With the results of the experiments I show that the recovery of elided structure starts at around 300 ms after onset of the critical word and is reflected by positive deflections. I argue that retrieval processes are underpinned by both linguistic and domain-general processes. In addition, I find a secondary positive component which I relate to more complex integration processes.

Chapter 7 investigates the difference between determiner *de* "the" and quantifiers *elke/alle* "every/all" in Stripping constructions. Again, I argue that retrieval processes are not exclusively steered by a syntax-related mechanism.

The experiment on prosody is reported in Chapter 8. I test the extent to which the prosody of the first conjunct predicts upcoming (deleted) structure. In an exploratory analysis, I show ERP effects related to attention/selection processes that are involved during the resolution of Gapping.

The overall findings, limitations and future prospects are discussed in the concluding Chapter 9.

CHAPTER 2

Theoretical background of Gapping and Stripping

In this chapter I discuss the distributional properties of Gapping and Stripping. The theoretical accounts that I review can be broadly categorised as syntax-oriented and semantics-oriented. Although this rather simplistic differentiation between syntactic and semantic accounts has been the driving force behind the current project, I argue that Gapping-like constructions cannot be captured in either syntactic or semantic terms. I further discuss an additional level of analysis, namely, prosody.

2.1 Objectives of theoretical ellipsis research

This chapter examines recurring topics in ellipsis research that I introduced in the previous chapter:

- What are the distributional properties of ellipsis: under which conditions is ellipsis permitted?
- How are elided phrases recovered? Which strategies (syntactic, semantic and prosodic-based) are involved?
- What is the division of labour between syntactic, semantic and prosodic-based mechanisms?

As the current study revolves around Gapping and its sub-type Stripping, this chapter defines Gapping-like constructions and explains how they are accounted for in the theoretical literature. Any theory on ellipsis should account for distributional properties – defining under which conditions ellipsis is permitted. Therefore Section 2.2 lists these conditions for Gapping and Stripping. As the differentiation between syntactic and semantic strategies is related to the distinction between "surface" and "deep" ellipsis types, I will take these notions into account in section 2.3 before exploring the different theoretical approaches in section 2.4. The different theoretical approaches will appear to be categorised on the basis of their point of departure: syntactic, semantic, and mixed. Crucially, the starting point pertains to the linguistic representation of the ellipsis site and the representation of the antecedent. In particular, approaches can be differentiated on their treatment of three crucial issues:

- What is the nature of the ellipsis site (i.e. its formal representation)?
- What is the nature of the antecedent (referred to as the "identity" condition)?
- Under which conditions is ellipsis allowed (referred to as the "licensing" condition)?

With respect to the ellipsis site, an ongoing debate concerns the question whether there is unpronounced syntactic structure. Related to this question are two restrictions on ellipsis, namely identity and licensing. Both terms became fashionable in the Generative literature since the 1990s with the publication of Lobeck (1995)'s book and subsequently took a central place in seminal works on ellipsis by Merchant (2001) and Aelbrecht (2010). An antecedent should be identifiable, i.e. recoverable. Identification of the antecedent is subject to some kind of parallelism, but the question is whether the antecedent's relation with the ellipsis site is constituted by means of syntactic or semantic terms. If a theory does not assume structure in the ellipsis site, the identity of the antecedent is by definition non-syntactic. However, structural accounts may differ as to the identity issue.

The notion of parallelism has been entertained not only syntactically and semantically, but also in terms of prosody. The last section of this chapter is dedicated to the prosodic aspect. Whichever theory of Gapping is preferred, it should be able to account for the distributional properties explaining under which conditions Gapping is licensed – finding a proper balance between syntactic, semantic and prosodic factors.

The variety of theories is almost endless. I will be concentrating on accounts that capture Gapping and thereby Stripping, incorporating different perspectives, in order to arrive at a well-balanced – albeit not exhaustive – overview of the literature. Evaluating the two sets of questions posed above, it will become clear why I use Gapping-like constructions in particular to probe ellipsis. Categorised as a highly-constrained surface ellipsis type, such constructions may be used to manipulate syntactic, semantic and prosodic complexity straightforwardly. But let us first observe the distributional properties of Gapping.

2.2 Diagnosing Gapping and Stripping

As discussed in Chapter 1.1, Gapping is characterised by an omission of at least the finite verb in the second conjunct of a coordinate structure as we can see in (1).

(1) a. De man kocht een boek, en de vrouw een krant.
the man buy.3SG.PST a book and the woman a newspaper
'The man bought a book, and the woman a newspaper.'

b. De mannen kochten een boek, maar de vrouw een krant.
the man buy.3PL.PST a book but the woman a newspaper
'The men bought a book, but the woman a newspaper.'

In Dutch Gapping constructions, apart from the connective *en* ('and'), *of* ('or') and *maar* ('but') may be used, though *of* and *maar* are not very common, as reported in a corpus study by Hoeksema (2007). The elided elements are semantically identical to their (linguistic) antecedents (see Hankamer & Sag, 1976; Neijt, 1979; Wyngaerd, 2007:2). However, it is not necessary that an elided verb has identical person, number, and gender features to the antecedent verb (see Repp, 2009:8-9).

The phrases in the right conjunct (*de vrouw* 'the woman' and *een krant* 'a newspaper') that contrast with their correlates in the left conjunct are called remnants. Typically, the remnants do not form a syntactic constituent. Kuno (1976) was probably the first to note that remnants must occur in a contrastive relation to their correlates. For example, *vrouw* versus *man* in (1a). Typically, non-clause-final correlates and remnants bear a rising pitch accent and clause-final correlates and remnants bear falling accents. The conjuncts themselves are separated by an intonational phrase break (Repp, 2009:14). The verb and other material that is elided in the right conjunct is "deaccented" in the first conjunct.

Example (2a) shows that, in tandem with the verb, an object may be elided. Additionally, (2b) shows that multiple remnants are possible, at least in Dutch (this possibility may differ between Dutch and English, see Jackendoff (1971)). Small capital letters indicate accented words.

(2) a. De MAN kocht een boek in LONDEN, en de VROUW in LEIDEN.
the man buy.3SG.PST a book in London and the woman in Leiden
'The man bought a book in London, and the woman in Leiden'

b. De MAN kocht een BOEK in LONDEN, en de VROUW een KRANT in LEIDEN.
the man buy.3SG.PST a book in London and the woman a newspaper in Leiden
'The man bought a book in London, and the woman a newspaper in Leiden.'

In both (2a) and (2b), the deaccented phrases in the left conjunct are taken into consideration to successfully interpret the right conjunct. Crucially, without drawing on these phrases, the right conjuncts are incomprehensible. In (2a), the phrase *kocht een boek* 'bought a book'and in (2b) *kocht* 'bought'are recovered and integrated with the remnants which yields a successful interpretation.

In principle, an adjunct contained in a Gapping construction is optional, in that it is not obligatorily incorporated in the interpretation; though Coppen, Borght, Dreumel, Oltmans, and Teunissen (1993) note that "adjuncts in the first conjunct that do not have a contrasting element in the second conjunct, are almost always filled in there" as is the case in (3).

(3) De MAN kocht een BOEK in Londen, en de VROUW een KRANT.
the man buy.3SG.PST a book in London and the woman a newspaper
'The man bought a book in London, and the woman a newspaper.'

Here, the interpreter automatically assumes *in Londen* to be the location of the woman buying a newspaper. Note that what is omitted – and interpreted – in the second conjunct does not necessarily form a constituent, as in the case of *kocht ... in Londen* 'bought in London'. This is another key property of Gapping (see for more examples Boone, 2014:21).

Negation, which can be seen as an adjunct in Dutch, is difficult to interpret in a gapped clause. In other words, it cannot be "filled in there". Whereas *in Londen* is recovered as adjunct in the Gapping construction in (3), this does not hold for *nooit* (4) In the examples in this chapter, I adopt conventional notation for indicating the acceptability of sentences, i.e. an asterisk * signals an ungrammatical sentence, while a single question mark indicates that the interpretation may be problematic. The notation "??" indicates that the interpretation is very problematic.

(4) ??De man kocht nooit een boek in Londen, en de vrouw een
the man buy.3SG.PST never a book in London and the woman a
krant.
newspaper
'The man never bought a book in London, and the woman a newspaper.'

I refer the interested reader to a study on negation and Gapping (Repp, 2009). For the time being, I would like to note that Repp's general conclusion is that negation in Gapping constructions can sometimes be interpreted if a proper combination of syntactic, semantic, discourse-pragmatic, and prosodic factors apply.

Hankamer and Sag (1976) have pointed out that as long as there exists a linguistic antecedent, Gapping can occur across a speaker boundary, i.e. with a different speaker producing the second conjunct. The assumption is that in both (5) and (6), speaker *Crit* has to recover the antecedents from a discourse level.

(5) a. Lisa: De man kocht een boek in Londen.
Lisa: the man buy.3SG.PST a book in London
Lisa: 'The man bought a book in London.'

b. Crit: En de vrouw een krant.
Crit: and the woman a newspaper
Crit: 'And the woman a newspaper.'

(6) a. Lisa: Wie gaat er mee een boek kopen?
Lisa who goes there with a book buy.INF
Lisa: 'Who is coming along to buy a book?'

b. Crit: Ik niet.
Crit I not
Crit: 'I am not'

Gapping occurs most often in coordinations connected by *en* ('and'), meanwhile it is disallowed in subordination. This is shown in (7a). Also, an antecedent that is embedded can only be omitted if the clause containing the gap conjoins with the embedded clause as we see in (7b) (Johnson, 2009).

(7) a. *De man kocht een boek, omdat de vrouw een krant.
the man buy.3SG.PST a book because the woman a newspaper
'The man bought a book, because the woman a newspaper.'

b. ?De jongen zei dat de man een boek kocht, en de vrouw een krant.
the boy say.3SG.PST that the man buy.3SG.PST a book and the woman a newspaper
'The boy said that the man bought a book, and the woman a newspaper.'

(7b) is ungrammatical if the ellipsis is interpreted as conjoining with *De jongen zei*. If it is interpreted as conjoined with the embedded clause, it is grammatical (see also footnote 6 in Johnson, 2009).

On the basis of the impossibility of subordination as seen in (7) and the fact that Gapping occurs in coordinate structures, Boone (2014:11) states that any account of Gapping should implement a restriction capturing the fact that "Gapping only occurs in coordinations where gap and antecedent are directly conjoined". He terms this restriction "Equal Conjunct Requirement". In slightly different terms, Winkler (2005:157) argues that Gapping is possible as long as the "Smallest Conjunct Constraint" is not violated. She further states that elliptical constructions with only one remnant together with an additive marker such as *ook* 'too'(as in (8)) bear a strong resemblance to the distributional properties of Gapping (Winkler, 2005:153-166).

(8) De man kocht vrijdag een boek in Londen, en de vrouw ook.
the man buy.3SG.PST Friday a book in London and the woman too
'The man bought a book in London on Friday, and the woman too.'

The elided structure *de vrouw ook* consists of one remnant (*de vrouw*) and an additive marker (*ook*). Such a construction is called "Stripping" and since it shares the relevant distributional properties of Gapping, I follow Hankamer & Sag, 1976 and Boone, 2014:10 in considering it a sub-type of Gapping.

With just this handful of examples, we have shed light not only on the distributional properties of Gapping, but also the issues that any theory of Gapping and Stripping must be able to account for. Let us now investigate how Gapping-like constructions have been entertained in the theoretical literature.

2.3 Surface versus deep

I have noted in the introduction that for interpretation of Gapping-like constructions a linguistic antecedent is needed. In the literature, Gapping (and other ellipsis types) have been likened to anaphora – items that refer to antecedents in order to recover meaning. Hankamer and Sag (1976) have suggested that elliptical structures may be classified as below, drawing on the distinction between surface and deep anaphora:

- Surface-ellipsis
  - on a par with surface anaphors (e.g. *himself*) which are bound to a linguistic antecedent
  - structural parallelism is a requirement
  - interpretation is done by reconstruction of the syntactic structure
- Deep-ellipsis
  - on a par with deep anaphors such as pronouns (e.g. *he*)
  - interpretation through referral to a conceptual level of representation
  - interpretation by rules of semantic interpretation

This differentiation assumes that surface-ellipsis requires syntactic parallelism between the elliptical structure and its antecedent, while for deep-ellipsis an "interpretive" approach holds. Deep anaphors "are not syntactically derived from full underlying forms" (Hankamer & Sag, 1976:423). According to the authors, surface anaphora must be consistent at a surface level with the anaphoric clause. This proposal is embedded in the context of Transformational Grammar (a precursor of the Minimalist Program) in which sentences are assumed to be transformed (derived) from a deep structure to a surface structure. Replacing the anaphor with the antecedent should yield a grammatical sentence. A surface anaphor is hypothesised to require "superficial syntactic identity of structure between the antecedent segment and the segment to be anaphorized" (1976:423), in contrast to a deep anaphor that is "not derived transformationally but is present in underlying representations" (1976:421) representing a semantic unit. As we will see, it is the distinction between surface and deep that has fuelled the difference between syntactic and semantic approaches to ellipsis. Therefore, we will adopt this classification as guidance for the remainder of this chapter.

Just as Gapping and Stripping are regarded as surface anaphora, Verb Phrase Ellipsis (VPE) is considered a surface anaphor as well. (9a) is an example of VPE: the VP *read the newspaper* is elided and replaced by *did too*. Note, that VPE contrasts with Stripping to the extent that VPE preserves a finite verb. The presence of the auxiliary *did* makes this an example of VPE and not a Stripping example.

(9) a. Pat read the newspaper, and Fran did too.
b. The newspaper was read by Pat, and Lee did it too.

[Murphy (1985b)]

According to Hankamer and Sag (1976) the deep anaphor *it* in (9b)[1] is understood at a "presyntactic" pragmatic level. It is "a semantic unit that appears elsewhere in the discourse or in context" (1976:420). Note, that although *did it too* in (9b) is closely related to VP-ellipsis *did too* in (9a), an anaphoric distinction is assumed. In their modified approach, Sag and Hankamer (1984) return to this issue suggesting that the distinction between deep and surface would mean that two different processing mechanisms are in charge, leading to peculiar situations to the extent that (10a) and (10b) would be processed totally differently.

(10) She told me to take the oats down to the bin,
a. so I did. (surface)
b. so I did it. (deep)

[Sag and Hankamer (1984)]

Sag and Hankamer (1984) propose that during sentence comprehension, the listener/reader builds two distinct representations, which are constructed in parallel, namely, "propositional representations (of the sentences of the immediately prior discourse) and discourse models (of the broader discourse context)" (1984:341). Since ellipsis "must be sensitive to scope of logical operators and variable binding", interpretation is assumed to be "determined by a propositional representation of the kind generally called *logical form*" (1984:328) – a representation that is not "surface" at all (but still a linguistic representation). They assume that all 'formerly known as' surface ellipsis types such as VPE, Sluicing, Gapping and Stripping are interpreted in terms of propositional representations. That is, "the interpretation of ellipses remains a rather simple copying of logical form". Additionally, Sag and Hankamer (1984)'s proposal can be used to distinguish general underspecification like the example in (4) (repeated here in (11)), from ellipsis (with linguistic antecedents), in terms of discourse grammar versus logical form.

(11) She waited there but he didn't show up.

While the approach of Hankamer and Sag (1976) sets the stage for comparing different theoretical approaches to the resolution of Gapping in the following sections, it is interesting that their paper analyses anaphoric relations in terms of "processes" rather than derivations. Apparently, they were working towards a bridge between competence and performance, which led to a model presented in 1984. This model can be regarded as 'performance' oriented at the

[1] I use the examples as given by Murphy (1985b) in review of Hankamer and Sag (1976)'s paper. The examples most effectively show the resemblance with Stripping.

core, yet driven by logical form as linguistic representation.

2.4 Approaches to the representation and derivation of ellipsis

Approaches to the representation and derivation of ellipsis may be distinguished as to the degree of syntactic structure assumed in the ellipsis site. As we will see, syntactic accounts argue for internal structure, while semantic accounts generally put more emphasis on the interpretation of elided structures. In addition, syntax-oriented accounts may consider the identity of a possible antecedent for ellipsis to be a fully-fledged syntactic structure. Let us first have a look at syntactic accounts of ellipsis that have focused on Gapping and Stripping.

2.4.1 Syntax-oriented accounts

Syntax-oriented accounts following the Chomskyan tradition assume that there is unpronounced structure in ellipsis sites at some point during the derivation. This is mainly driven by the observation that Gapping leaves non-constituents at surface structure. Assuming unpronounced structure would help to reconcile the syntactic status of remnants. With respect to accounts of Gapping three topics are central and every account utilises one or more of the following operations:

- Deletion: at a relatively late point in a derivation some structure is elided at PF.
- Movement: at some point in a derivation constituents move to take scope over a (later) elided structure, or they move rightwards, or they move without elision establishing a dependency relation between antecedent and gap.
- Copying or sharing: at some point in a derivation some structure is copied or shared at PF/LF.

The informed reader may notice that anaphoric accounts such as Fiengo and May (1994), nowadays referred to as proform theories, are not listed. Syntax-oriented accounts which assume that throughout a derivation the ellipsis (VPE in particular) is treated as null-element or anaphor which can be linked to an antecedent structure, usually understand Gapping as a separate form of ellipsis – as surface instead of deep anaphor (c.f. Lobeck, 1995; Williams, 1977). Generally, such accounts strive to characterise symmetrical (or parallel) anaphoric relations. For example, Fiengo and May (1994:83) propose that "what is structurally represented are indices, which by hypothesis are complex objects, consisting of a value and a type, and it is indices and their relations that ultimately have semantic import." As we will see in section 2.4.2, there is a possibility to incorporate a representation of elided phrases in mere semantic terms, but theories that do so put less burden on syntactic structure. Let us first start with the first theory of Gapping.

Deletion: the emergence of Gapping research

In the generative literature, Ross' dissertation (1967) is regarded as the first attempt to account for conjunction ellipsis. He proposes that the derivation of Gapping constructions is subject to a "Conjunction Reduction Rule" (Ross, 1967:100), which should be taken as a general assumed rule within the transformational framework (a precursor of Minimalism). For example, such a rule would also reduce sentences as (12a) and (13a):

(12) a. John knows the answer and Bill knows the answer.
b. John and Bill know the answer.

(13) a. Otto sells Buicks and Otto sells Fords.
b. Otto sells Buicks and Fords.

[Ross (1967:116)]

In the b-examples above, Ross assumes that – underlyingly – a full structure is available. Elaborating on conjunctions, Ross (1967) proposes the "Coordinate Structure Constraint" as defined in (14).

(14) **The Coordinate Structure Constraint**
In a coordinate structure, no conjunct may be moved, nor any element contained in a conjunct be moved out of that conjunct.

[Ross (1967:161)]

Ross (1967) notes that (14) is needed since a conjoined NP cannot be questioned. In the transformational framework and its successors, movement is assumed to be a core operation to linearise phrases while keeping track of the original positions for interpretation purposes (among other things). It is a way of establishing a dependency relation. *What sofa* in (15) has been moved out of the conjunction, which is impossible as no dependency between the source and the moved element can be established.

(15) *What sofa will he put the chair between some table and?

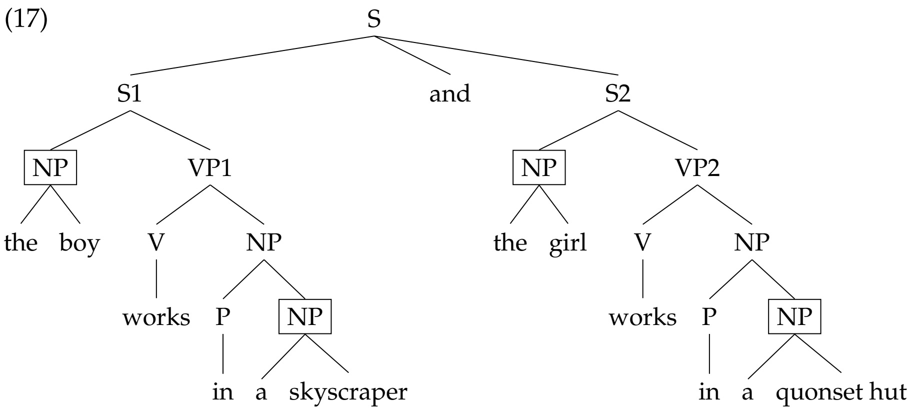
[Ross (1967:158)]

Further, Ross (1967:171) suggests that (14) could "provide a test for coordinate structure" to the extent that Gapping constructions *are* coordinate structures. As a consequence, the underlying structure of a Gapping construction should reflect this, i.e. two conjoined sentence nodes. In (16b) we see Ross' example of Gapping, which he assumes to have an underlying structure as (16a) – represented in a tree structure (17).

(16) a. The boy works in a skyscraper and the girl works in a quonset hut.
b. The boy works in a skyscraper and the girl in a quonset hut.

[Ross (1967:171)]

(17)



[Ross (1967:171); I follow him in projecting prepositional phrases as NPs.]

Ross evaluates the underlying structure of (16b) as is pictured in (17). He assumes the second conjunct to be a full bodied sentence structure: in his terms the S2 node is not "pruned", contrasting the VP2 node that *can* be pruned in order to attach the NP containing *in a quonset hut* to the S2 node. This would still yield a coordination of two clauses. Ross further argues that if S2 is pruned, only one clause would remain in which the boxed NPs should be movable. However, this is not the case as is seen in (18).

(18) a. *Which boy works in a skyscraper and the girl works in a quonset hut?
b. *The skyscraper which the boy works in and the girl in a quonset hut belongs to Uncle Sam
c. *The girl who the boy works in a skyscraper and in a quonset hut has a dimple on her nose
d. *Which quonset hut does the boy work in a skyscraper and the girl in?

[Ross (1967:172)]

Since movement of NPs of either conjunct leads to ungrammaticality, Ross suggests that both conjuncts have the same (parallel) underlying syntactic structure – in other words they are two conjoined sentences.

Although frameworks have changed throughout the years, ever since Ross' work scholars have been continuing to put forward proposals based on his ideas. In the current framework, Minimalism, the level of PF is generally assumed to be the point at which deletion takes place. It should be noted that from the very beginning, together with the syntax-centred approach, additional licensing constraints have been put forward (although the term "licensing" would be coined later). For example, Sag (1976) and Neijt (1979) already considered semantic identity of antecedent and gap necessary. This idea has stood the test of time and can be found in adapted forms in Merchant (2001), an influential work on Sluicing and ellipsis in general, and in Coppock (2001), who builds on this in combination with movement. Hartmann (2000) is an example of combining prosodic elements in a deletion account, a suggestion already made by Sag (1976:294-295) while only touching on this topic briefly.

Movement and deletion

Recently, Boone (2014) offered an account in which Gapping, Stripping and Fragment Answers are derived by movement of remnants followed by deletion. Building on earlier movements accounts (such as proposed by Aelbrecht, 2007; Johnson, 2009; Merchant, 2005) he assumes that remnants "escape" the ellipsis site by moving leftward (higher up the structure) while remaining phrases are subject to deletion. The method is shown in (19a) and crucially, the movement is exceptional since it may not occur when ellipsis does not take place as in (19b). A "t" indicates a trace of a moved phrase.

(19) a. [$_{S1}$ The boy works in a skyscraper] and [$_{S2}$ [$_{NP}$ the girl]$_i$ [$_{PP}$ in a quonset hut]$_j$ [$_{XP}$ t$_i$ ~~works~~ t$_j$]]]]
b. *[$_{S1}$ The boy works in a skyscraper] and [$_{S2}$ [$_{NP}$ the girl]$_i$ [$_{PP}$ in a quonset hut]$_j$ [$_{XP}$ t$_i$ works t$_j$]]]]

Boone shows in his dissertation that movement of the remnants solves the issue of non-constituency of the elided structure. In (19) *XP* marks the constituent that may be targeted for ellipsis. The advantages of his approach is even more evident if the ellipsis is discontinuous as is shown in (20).

(20) a. The boy works in a skyscraper with great pleasure and the girl ~~works~~ in a quonset hut ~~with great pleasure~~.
b. [$_{S1}$ The boy works in a skyscraper with great pleasure] and [$_{S2}$ [$_{NP}$ the girl]$_i$ [$_{PP}$ in a quonset hut]$_j$ [$_{XP}$ t$_i$ ~~works~~ t$_j$ ~~with great pleasure~~]]]]

As licensing condition, Boone proposes a discourse-dependent constraint which checks that a non-hierarchical structure exists between two conjuncts. The proposal is very much inspired by the notion of D(iscourse)-linking originally proposed by Pesetsky (1987). While Pesetsky used this notion to argue for syntactic movement at the level of LF of non-D-linked *wh*-phrases, D-linking has been widely used by syntax-oriented scholars to accommodate discourse (or semantics) related representations. Boone takes seriously the symmetric discourse relation that exists between remnants and antecedents (as argued by Culicover & Jackendoff, 2005; Kehler, 2000; Levin & Prince,

1986) though he represents the discourse relation as syntactic trees. However, it is questionable to what extent a discourse relation fits in a Minimalistic model as depicted in Figure 1.2. Such a relation might be better abstracted in semantic terms. His approach is attractive in that "the licensing condition on ellipsis is not a condition specific to ellipsis, but follows from a general requirement on recoverability" (Boone, 2014:9). In other words, since no special stipulations for ellipsis need to be postulated, as a consequence, the essence of ellipsis processing may not be different from normal sentence processing.

Boone further states that:

> Theories that refrain from postulating syntactic structure in Gapping and Fragments must invoke mechanisms that ensure that the remnants of ellipsis have the same properties and show the same behavior as they do in the corresponding non-elliptical utterance. Although such mechanisms can no doubt be hypothesized, they unnecessarily complicate the grammar. If we accept that there is syntactic structure in the ellipsis site, the connectivity facts follow straightforwardly, without the need to postulate additional conditions and constraints.
>
> [Boone (2014:33)]

One could take issue with the notion of "complicating the grammar". Boone proposes that Gapping and Fragments are subject to movement and deletion procedures. A theory that does not assume these mechanisms as core operations may use other ones. This does not entail that an alternative theory complicates the grammar by definition. Further, he proposes the rather circular argument "the fact that remnants move out of the ellipsis site, constitutes additional evidence for structure in the ellipsis site, since extraction entails that there is syntactic structure to extract from" (2014:33). If one assumes structure, one could assume movement – in this order. If hidden levels of structure are not assumed, movement can never be acknowledged as "fact". Furthermore, his assumption concerning movement of the remnants is a case of "Exceptional Movement" which is only applicable to ellipsis. Quite easily, this specialised mechanism may be considered as an instance of complicating the grammar in its own right.

Copying

While Sag and Hankamer (1984) can be seen as a precursor of LF-copy accounts, a recent derivational copying account for Gapping is put forward by Repp (2009). Kobele (2015) has also proposed a general copying account of ellipsis derivation including Gapping. Let us look at them briefly in turn.

Repp (2009) proposes that in Gapping the second conjunct is derived by copying the remnants from the left conjunct that are spelt out at PF. The elided material in the right conjunct is only visible at LF. In that sense, the copying mechanism is comparable to an LF-copy account, as is proposed for Sluicing by Chung, Ladusaw, and McCloskey (1995). However, Repp implements the copying mechanism as "sideward movement" which means that copied material may be merged with unconnected independently-derived syntactic objects. Sideward movement could be seen as a repair mechanism for derivations that may crash (Repp, 2009:31) and occurs in two steps. The second conjunct is derived in parallel with the first conjunct and after the first conjunct has been sent off to PF, missing material may be copied. For example, an elided verb is copied from the left conjunct and is moved to the right conjunct and merged with the remnant object. If not, the "numeration" of the remnants would yield an incomplete sentence – a crash.

While a verb (lexical projections) and object (functional projection) may be copied, Repp argues that adjuncts are not "sentential functional projections" (2009:43) and can therefore not be copied. This is helpful to account for the behaviour of negation in Gapping as Repp shows in her book. However, we have seen that adjuncts in Dutch may be filled in (see section 2.2 above). Repp (2009:80) argues that adjuncts in Dutch are not obligatorily filled in and assumes that processes of "accommodation", mechanisms at the level of discourse and pragmatics, are at work in such cases. While the details of accommodation are lacking, it is not made clear why negation should be immune to accommodation. Also, a consequence of her approach would be that in Dutch sentences such as (3) repeated here in (21), two mechanisms of different levels of representation would be at work in parallel by default – sideward movement and accommodation. If the adjunct is taken into consideration for further interpretation, it is generally assumed by structural accounts that it attaches directly to the recovered VP node *kocht een krant*.

(21) De man kocht een boek in Londen, en de vrouw een
the man buy.3SG.PST a book in London and the woman a
krant.
newspaper
'The man bought a book in London, and the woman a newspaper.'

Kobele (2015) takes a more loose attitude as to the question of whether there is structure in the ellipsis site. Rather, he asks how much syntactic information is needed to account for structural requirements on grammatical ellipsis. Instead of reconstructing syntactic structure, possible antecedents should have the proper syntactic category. In that sense Kobele apprehends ellipsis as reuse of some syntactic structure rather than "recomputation". During the derivation, it would suffice to select the semantically appropriate antecedent from a possible set while syntax is mainly working as a "filter".

Not stated as such in his paper, Kobele's proposal contrasts with a sideward movement approach to accounting for repair operations. He further points to the fact that copying is related to deletion, that is, "copying theories should be thought of as descriptions of the algorithm implementing deletion theories". Furthermore, proposing a copying mechanism, although stated as derivation "procedure", still leaves us with the question of how it would be implemented as a parser. It is therefore impossible to distinguish between deletion (possibly in combination with leftward movement as Boone proposes) and copying accounts using psycholinguistic data, meaning that – from a processing perspective – we may lump deletion and copy accounts together.

Although Kobele's approach is syntactic in nature, contingent on Minimalist notions such as merge and move, he takes seriously the semantic relation that the antecedent and ellipsis site constitute. In contrast to the LF-copy account by Chung et al. (1995) (who regard the basic nature of LF as structural rather than semantic), ellipsis sites are thought to be resolved by replacing them with their antecedents semantically, while antecedents are delimited syntactically. In this sense, he is leaning towards semantic approaches using insights that we will consider in the next section.

2.4.2 Semantics-oriented accounts

Semantics-oriented accounts put less emphasis on syntax (though they do not neglect it) and are therefore autonomous from a strictly derivational model as shown in Figure 1.2, but can be related to the tripartite model as presented in Figure 1.1. After introducing a "matching" condition, I will briefly discuss a related discourse condition, a Head-driven Phrase Structure Grammar approach and a Categorial Grammar solution. I will end with a proposal within the framework of Simpler Syntax that may be understood as a refined version of an integrated matching and discourse condition.

Matching condition

Although the importance of a semantic match between antecedent and ellipsis was already known in the early days of ellipsis research, Dalrymple, Shieber, and Pereira (1991)'s paper marks the beginning of approaches that emphasise the semantic import during resolution, moving syntax into the background. Using lambda abstraction (i.e. symbolic expressions of semantic abstraction), they propose a higher unification algorithm to be in charge of the recovery of the ellipsis antecedent. Unification is an algorithmic process – used in logic and computer science – to solve equations between symbolic expressions. Notably, this process goes beyond the (structural) representation at LF that is assumed in syntax-oriented accounts. At the same time, the authors argue that their proposed process maps easily onto a discourse model. Therefore, CS in the tripartite framework in Figure 1.1 seems the appropriate level at which this process may take place.

The clause *Fran likes cheese* in (22a) can be decomposed into an underspecified property *P*. When this property is applied to the interpretation of the subject (*Fran*), it will yield the interpretation of the clause. (22b) shows how the property maps to a lambda expression in which the subject is a variable. Predicating the property of the subject *Fran* as is done in (22c) will yield the interpretation of (22a). Quite often, the reduced denotation as in (22d) is used to show the end product. Note that these denotations assume a configuration in which a property is applied to the interpretation of the *subject* and not the *object*. This decision may hinge on syntactic information types.

(22) a. Fran likes cheese.
b. $P \mapsto \lambda x.like(x, cheese)$
c. $\lambda x.like(x, cheese)(Fran)$
d. $like(Fran, cheese)$

Expanding this idea to ellipsis, as exemplified in the Stripping example in (23a), recovery can be a matter of finding the matching property shown in (23b), which can be predicated of the subjects accordingly to get the interpretation of both the left and the right conjunct. The simplified solution to the equation of the lambda expression is shown in (23c) giving the interpretation of both conjuncts.

(23) a. Fran likes cheese, and Leo too.
b. $\lambda x.like(x, cheese)$
c. $like(Fran, cheese) \wedge like(Leo, cheese)$

In sum, Dalrymple et al. assume "a connection between the syntactic and the semantic representation of the source sentence" to guarantee that "solutions produced by higher-order unification satisfy the constraint that parallelism must be maintained by abstracting out of parallel positions" (1991:406-407). In this sense, this proposal acknowledges the syntactic structure in terms of relation formation. However, the authors distance themselves from identity-of-relations analyses that assume that the interpretation of ellipsis " involve[s] copying the interpretation of a constituent in the source" (1991:437); rather, the predicate argument relation is extracted from the source clause. As a consequence, their approach does not need additional syntactic machinery for quantifiers such as *every* and *all*. For example, in (24b) it is shown how the shared property of (24a) would be represented. By contrast, syntactic accounts usually represent quantified phrases at LF yielding extended tree structures. Only after "Quantifier Raising" (QR) – a form of movement – has taken place may the structure be sent off for further interpretation. This implies a burden on mechanisms of movement and/or copying since extra structural information has to be analysed. An LF structure for the first conjunct of (24a) is given in (24c). For the sake of contrast, an LF representation of the utterance in (22) *Fran likes cheese* without a quantified object is shown in (24d).

(24) a. Fran likes every cheese, and Leo too.
b. $\lambda z.every(x, cheese(x), like(z, x))$
c. $[_{IP}\ [_{DP}\ \text{Fran}]_1\ [_{VP}\ \text{every cheese}]_2\ [_{VP}\ t_1\ \text{loves}\ t_2]]]]$
d. $[_{IP}\ [_{DP}\ \text{Fran}]_1\ [_{VP}\ t_1\ \text{loves cheese}]]]$

The lambda term of the utterance containing a quantified phrase is slightly more complicated in (24b) compared to the one in (23b), but no additional semantic stipulations are required. Note again, however, that (24b) implies a certain syntactic configuration. Matching of the property may apply to resolve the ellipsis. Boone (2014:135-137) explains that QR is problematic for a Minimalistic approach concluding that it is a form of movement that falls outside the computation before Spell-Out (see Figure 1.2). Just as the exceptional movement he promotes to arrive at an appropriate string at PF for the elliptic structure, quantified phrases are subject to an additional exceptional movement to yield the appropriate LF structure at the ellipsis site. In a footnote, Dalrymple et al. emphasise that their analysis concerns:

> its use of an equational framework for declaratively characterizing ellipsis resolution, not its use of particular logics for the representation of meanings. Nonetheless, the use of typed lambda calculus allows us to directly state our analysis with a minimum of extraneous machinery.
>
> [Dalrymple et al. (1991:fn 9)]

In other words, a syntactic representation at LF may be hypothesised but it is not a necessary requirement for the resolution mechanism to work.

While their approach has been very influential, it has also met some criticism as the proposed theory would "overgenerate", which means that it would allow for constructions that are not acceptable. This contrasts with syntax-oriented accounts that generally "undergenerate" if they ignore other levels of representation. In terms of implementation Dalrymple et al.'s advantage is that they put forward a theory-neutral account which can be used in different corners of linguistic theory.

The authors, however, do not show how adjuncts of the kind as shown in (25) should be handled. For this dissertation it will transpire that it is necessary to have a better understanding of adjuncts, given that they will be integral to the design of one of the experiments reported in later chapters.

(25) Fran likes cheese in the morning, and Leo chess.

There are two reasons to assume that a semantic denotation of the first clause in (25) could be expressed by means of a conjunction of two propositions as is represented in (26).

(26) $like(Fran, cheese) \wedge in(the\ morning)$

First of all, the conjunction predicts the proper entailment, since (25) entails that *Fran likes cheese* and that some action happens *in the morning*. If the first proposition is embedded in the adjunct proposition it entails that liking of cheese is dependent on the time of day. This is, however, not the case.

Secondly, adjuncts, especially those denoting time or place, may be "filled in" in elliptical constructions, and in Dutch they usually are, as was demonstrated in (3) above. This would exclude the possibility of embedding the adjunct in the first proposition. However, following Dalrymple et al.'s account, we could assume that the interpreter matches the predicate argument relation between antecedent and remnant(s) while the interpretative conjunction allows for a choice as to the extent to which an elided adjunct – a proposition that is readily available – should be conjoined to the recovered proposition. Since this operation is semantic in nature, interaction with processes of accommodation to account for optionality is less unexpected than is the case in a structural account. Jackendoff (2007) would agree with such an approach since he suggests that "there is very little syntactic constraint on either the form of the adverbial or its position, and the semantics and pragmatics are doing most of the work."

On a par with Repp (2009), a semantic approach such as Dalrymple et al.'s classifies negation as a special kind of linguistic unit; not as a movable adjunct, however, but as an "operator". Assuming that negation as an operator can take wide scope over the propositions as in (27), this would yield a negative state of the whole proposition. In other words, negation can be sentential.

(27) $\neg[like(Fran, cheese) \wedge in(the\ morning)]$

This may give the impression that an elliptical construction needs reference to a proposition that is true. Possibly, an additional semantic constraint may be that Gapping constructions can only be linked to a state of affairs that exists. I will leave this issue for future research noting that a semantic approach predicts quite naturally that adjunct-hood in general is not exceptional by default, but negation is.

Dalrymple et al. suggest that the semantically represented operator "neg" has a parallel operator "pos" in the right conjunct. They use the VP-ellipsis example in (28a) to show this. The property of the left conjunct is represented in (28b) in which $\lambda S.S$ refers to the underspecified operator that renders the proposition either positive or negative.

(28) a. Dan didn't leave, but George did.
b. $P \mapsto \lambda x.\lambda S.S(left(x))$

In elliptical sentences, Dalrymple et al. assume $\lambda S.S$ to be a property applied to the interpretation of "neg" in the left conjunct since there is a negation. The right conjunct delivers a positive operator which is derived from the affirmative property of *did* (and maybe the contrastive *but* plays a role here too). Therefore, the negation has not been elided but contrasted. If no additional operator is available in the right conjunct, it might be that a parallelism condition on operators holds, meaning that there should be an overt operator of some kind, or else the operator of the left conjunct applies. Gapping constructions with negation in the left clause can be saved by adding an operator in the right clause – regardless of the polarity – as is demonstrated in (29).

(29) De man kocht niet een boek, en de vrouw *(wel/niet) een
the man buy.3SG.PST not a book and the woman (AFF/NEG) a
krant.
newspaper
'The man didn't buy a book, and the woman (did/didn't) a newspaper.'

If no overt operator is available in the right conjunct, it seems that the predicate argument relation that is abstracted from the left clause is assigned the interpretation to the operator that is available, yielding a negative state that can not be interpreted as such. Again, comparable to Repp's account, processes of accommodation may steer the interpretation of the sentential variable $\lambda S.S$ in some cases. The bottom line, still, is that negation seems to be the exception – not adjunct-hood in general. Crucial implication in favour of Dalrymple et al.'s account is that processes of accommodation at the level of CS target semantic representations more easily than they would under Repp's account as these processes would target representations at a different level, that would be impenetrable under minimalistic assumptions. Note though, that Dalrymple et al.'s account may need to be constrained by syntactic information types, for example, to prevent it from overgenerating.

Discourse condition

Hardt (1993:41) proposes a formal process in which meanings are assumed to be stored in a discourse model, while anaphoric expressions are assumed not to have predetermined antecedents. Rather, they are determined at some stage during the derivation. This computational model is apt to resolve VP-ellipsis, while it analyses elided structures as "proform" (also called "null-proform") or "proverb" (on a par with anaphors). Governed by a semantic identity condition, a missing VP is treated as a variable that is semantically interpreted just as other variables such as pronouns.

Together with Dalrymple et al., Hardt's approach falls under the so-called proform theories of ellipsis. Generally, such theories, starting with Wasow (1972) and later developed by proponents such as Fiengo and May (1994) and Lobeck (1995), assume that deep ellipsis forms should be treated as anaphora assuming that the ellipsis site is a null-pronoun (but see Baltin, 2012 for an argument that proforms involve deletion). Proform accounts assume that antecedents are semantic objects, and that ellipsis sites are resolved by replacing them with their antecedent in the semantics. In contrast to syntactic proform theories, Dalrymple et al. and Hardt do not assume syntactic structure at the ellipsis site. Hardt proposes that the interpretation of proforms is accomplished through store and retrieve operations that make reference to a discourse representation. At the same time, neither Gapping nor Stripping have been taken as serious candidates to be accounted for by means of a proform method. Likewise, Hardt (1993:122) argues that Gapping does not "require access to the discourse model for recovery of missing material." A distinction is made between proforms (VP-ellipsis) and conjunction forms (Gapping, PseudoGapping, Stripping). Both forms are treated by Dalrymple et al. (1991) in the same vein. Hardt (1993:123), however, suggests that their method is too unconstrained ("overgenerates") since the parser would deal with ungrammatical sentences effortlessly. Therefore, Hardt suggests that the focus based mechanism such as proposed by Rooth (1992) is an alternative interpreter for conjunction forms. Following this theory, focus helps to determine the relation between antecedent and remnant(s) at a semantic level of representation. Still, the question remains whether two separate mechanisms are preferred. To some extent, Hardt's approach is a first move towards to a Simpler Syntax approach, which I will discuss after an excursion to HPSG and Categorial Grammar approaches.

A HPSG approach

Head-driven Phrase Structure Grammar (HPSG) is a unification-based non-transformational theory. It is a constraint-based theory, which means that constraints license small pieces of linguistic structure. It utilises multiple inheritance type hierarchies and unification of typed feature structures as central formal mechanisms (see Pollard and Sag, 1994 for an in-depth explanation). An utterance in HPSG is represented as feature structures of type *sign*. For example a sign can be of type *word* which is stored in the lexicon as an entry containing descriptions of (or constraints on) feature structures. Type *phrase* will contain phrase structure rules, construction rules or "immediate dominance schemata". A notable aspect of HPSG is that phonological, syntactic, semantic and contextual features may be obtained at once, fitting the parallel architecture as such as shown in Figure 1.1.

An HPSG account for Sluicing and Fragments, elliptical constructions which are subject to a resolution strategy that is applicable to Gapping (as we have seen in Boone, 2014), has been proposed by Ginzburg and Sag (2000). They propose that Fragments are introduced by the phrasal-type headed fragment-phrase which must dominate the Fragment. The type *phrase* contains the type *local* which specifies the values for CATEGORY and CONTENT.[2] Their solution is to posit that these values are the same as those of the correlate – without assuming any syntactic structure in the ellipsis site. This is not to say that every HPSG account abstains from assuming syntactic structure. For example, Lappin (1999) proposes a mechanism based on syntactic reconstruction.

Ginzburg and Sag 2000:301 take it that prior syntactic structure only functions to determine the appropriateness of the utterance. At the same time, contextual information is available based on the notion "question under discussion" (QUD) – a set consisting of the currently discussable questions among

[2]I follow the convention that values for types in HPSG are written in capitals.

dialogue participants. For example, in (22a), here repeated as (30a), the implicit QUD would be "Who likes what?".

(30) a. Fran likes cheese.
b. Fran likes cheese, and Leo chess.

Ellipsis is licensed when the current QUD is answered by the remnants. This QUD can be overt, as in Fragments that answer overt questions, but the QUD may also be implicit, as is the case in Gapping constructions as in (30b), in which *Leo chess* answers the QUD "Who likes what?". Ginzburg and Sag (2000) suggest that the implicit QUD could be represented as CONTEXT feature containing a parameter that sets the list of open propositions (among other things), which are solved by the remnants. To a certain extent, this reminds us of the underspecified property used by Dalrymple et al. (1991), though Ginzburg and Sag (2000:298) allow for the possibility to incorporate sufficient syntactic sensitivity in their account to overcome overgeneration, that is, to develop a syntactic filter. Further, as we will see in section 2.5, a QUD approach is closely related to the way focus is assigned.

A Categorial Grammar approach

In Combinatory Categorial Grammar (CCG), which is a version of Categorial Grammar, an element like a verb is associated with a syntactic category – comparable to CATEGORY in HPSG mentioned earlier. This category identifies the verb as a function and specifies the type and directionality of the verb's argument(s) (depending on the transitivity) and the type of the result of the function. The effect is that a category can be understood as a combined syntactic-semantic object constituting a transparent interface between surface syntax of a sentence and its underlying semantic representation. Steedman (1990) has become known for his analysis of coordination, including Gapping constructions. Enabled by the flexible notion of constituency which is a core aspect of CCG, it exploits predicate-argument relations rather than syntactic structures. In fact, the only level of representation assumed is the predicate-argument structure. Rather than a module of grammar, "syntax" can be seen in CCG as a history of the algorithm that is used to determine the predicate-argument structure on the basis of a given surface string (during comprehension) – comparable to a "compiler" in certain computational approaches.

Steedman (1990:234) proposes that "Gapping requires the recovery of the arguments from the left conjunct, rather than the recovery of the verb". Treating Gapping as constituent coordination, he assumes that categories of the left conjunct can be decomposed. The 'remnant constituent' retrieves the arguments of the left conjunct upon which interpretation proceeds. Steedman (1990:255) states that Gapping resolution is "purely syntactic, and not to be mediated by anaphora of any kind, pragmatically specialized though it is". Note though, that CCG exploits logical form rather than the surface structure. The left conjunct decomposition operation resembles to some extent the higher-order unification method that we encountered in 2.4.2 (as is noted in Dalrymple et al., 1991:fn 3) and further compares with an interpretational method such as proposed by Cremers (1983). A common aspect of accounts that denote ellipses as sets of functions with propositional values is that they can account for the apparent lack of proper syntactic constituency in elliptical structures. Interestingly, such interpretational accounts are relatively easy to transpose to computational methods as we will see in section 2.6.

A Simpler Syntax approach

In the Simpler Syntax enterprise proposed by Culicover and Jackendoff (2005:273-282), Gapping constructions are assumed to be resolved in the syntax-semantics interface. Their theory pictures a situation in which a non-derivational syntax does not employ hidden levels of representation. Not surprisingly, it compares with constraint-based lexicalist theories such as HPSG and is compatible with the tripartite architecture in Figure 1.1. Just as Ginzburg and Sag (2000), Culicover and Jackendoff (2005) assume that language users interpret Gapping constructions without reconstructing a derivation from an underlying structure. In a sentence such as (31a), Culicover and Jackendoff treat remnants *Leo* and *cheese* as so called "orphans" – paired fragments that make up a constituent of no specific category and that are represented in the syntax. Again, we see flexibility in terms of syntactic constituency. In the semantics, the function $\mathcal{F}$ connects the orphans as arguments, a procedure the authors refer to as "Indirect Licensing"(IL) – a discourse-based mechanism. The missing material is recovered from the non-focused part of the source clause. This way, the interpretation and licensing of Gapping depend mainly on a Conceptual Structure function (see Figure 1.1) which "amounts to the presupposition of the antecedent, constructed by substituting variables for the two foci in the CS of the antecedent" (Culicover & Jackendoff, 2005:276). They reason that the remnants' contrastive focus hinges on the presupposition of the antecedent taking a position that is also apparent in semantics-based approaches. They are, however, less detailed as to the extent of the autonomy of and interaction with prosodic rules, as focus hinges entirely on CS. Presumably though, the phonological structure of the left conjunct should easily be employed considering the parallel nature of the framework. The syntax and semantics of the elliptical conjunct of (31a) are represented in (31b).

(31) a. Fran likes cheese, and $[_{XP}$ Leo$]$ $[_{YP}$ cheese$]$.
b. Syntax of the right conjunct: $[\mathrm{XP}_i^{ORPH_1}\ \mathrm{YP}_j^{ORPH_2}]$
CS of the right conjunct: $\mathcal{F}([\mathrm{X}_i$ C-FOCUS$]$, $[\mathrm{Y}_j$ C-FOCUS$])$

While the contrastive foci of the orphans (i.e. $[\mathrm{X}_i$ C-FOCUS$]$ and $[\mathrm{Y}_j$ C-FOCUS$]$) are used to match the contrasting foci of the antecedent structure, the function $\mathcal{F}$ in (31b) can be utilised to replace the corresponding foci of the first conjunct by variables which would yield a reduced lambda denotation such as was introduced earlier in this section: like(X, Y). Culicover and

Jackendoff subsequently assume that the contrastive foci of the right conjunct (as represented above) can be substituted giving the interpretation of this clause: [like(Leo, chess)]. The connection with propositional abstraction accounts like Dalrymple et al. (1991) is apparent. In a talk, Merchant (2007) suggested that Indirect Licensing can be seen as an updated version of Hardt (1993) and Dalrymple et al., but he remains critical as to the importance of structural issues. A recurring example is that of "connectivity effects". For example, in (32) it is shown that a Fragment Answer must carry the underlying voice of the question.

(32) a. Q: Who is ordering pizza? A: Pizza is being ordered by Fran.
b. Q: Who is ordering pizza? A: *By Fran.
c. Q: Who is ordering pizza? A: Fran.

Since the answer in (32b) cannot be uttered in passive voice – which is possible in principle as shown in (32a) – it is argued that an underlying syntactic structure should be assumed. (32c) is grammatical as the Fragment Answer follows the voice of the question. Culicover and Jackendoff (2005:539) suggest, however, that connectivity effects can be accounted for within their theory since IL transmits syntactic properties from the discourse antecedent to the Fragment. In conclusion, IL does not only exploit CS but capitalises on the parallel nature of the framework that this mechanism is embedded in. In addition, IL overcomes the criticism related to overgeneration of ungrammatical utterances in a purely semantic account; also it allows for an extension to account for other ellipsis types, resolving the disparity between deep and surface ellipsis types and the accompanying mechanisms to interpret them.

2.5 Parallelism, identity and focus

While reviewing the various accounts in the previous sections it has become clear that parallelism plays a considerable role. Syntax-oriented approaches emphasise the requirement of parallel syntactic structure between two conjuncts, and semantics-oriented accounts generally emphasise identity of relations. Some authors refer to "matching" rather than identity, which seems, following the definition in (33), more useful.

(33) To match: corresponding in some essential respect with something previously mentioned or chosen: a new coat and a hat to match.

[McKean (2005)]

Matching allows for using just enough information – for example, just enough syntactic information – to overcome connectivity effects. This contrasts with "isomorphism" which is also used for identity. In derivational theories, syntactic isomorphism extends to the lexical item and not so much to the inflection, which is assumed to be regulated by a higher node. Therefore, the ellipsis in (34a) is understood as in (34b) while being licensed under syntactic identity.

(34) a. These men are more clever than Mary.
b. These men **are** more clever than Mary **is** clever.

[Vicente (2008)]

Some structural accounts that speak of matching of 'just enough syntactic information' (see for example Kobele, 2015) and matching of truth conditions (which has been quite common since Sag, 1976), are comparable to semantics-oriented accounts in that they take pieces of semantic information of the antecedent into account. In general, semantic parallelism seems stronger than syntactic isomorphism as is pointed out in the following excerpt from Vicente (2008).

> [...] while both semantic parallelism and syntactic isomorphism play a role in the licensing of ellipsis, their relative importance is different. Semantic parallelism appears to be a truly inviolable condition; on the other hand syntactic isomorphism can be violated under certain specific conditions. The fact that these two conditions have different rules of application entails that they constitute separate conditions, neither one being reducible to the other.
>
> [Vicente (2008:21)]

In other words, semantic and syntactic constraints may apply in parallel, which entails that both semantic and syntactic representations are relevant levels in order to resolve ellipsis. Especially derivational accounts may encounter problems as long as they do not assume an interaction between semantic and syntactic representations. Crucial in Vicente (2008)'s analysis is the role of focus structure that can violate syntactic isomorphism. This brings us to the issue of prosodic parallelism.

Kuno (1976) precipitated a line of research with respect to prosodic parallelism – which one could call a syntax-discourse interface tradition – implementing focus to account for ellipsis phenomena. Kuno notes that the acceptability of Gapping is dependent on discourse context.

> [...] Gapping is a pattern that is used for presenting contrastive pairs of information segments, and [that], because of this semantic function, the constituents left over after Gapping must represent new, unpredictable information [...]
>
> [Kuno (1976:309)]

Kuno proposes that focus is driven by semantics and that in Gapping, the deleted elements must be given, which results in an unpronounced antecedent, a contrastive topic constituted by the first remnant and a contrastive focus on the second remnant.

Although Kuno's approach is "functional" at heart, which means that its starting point is language as a communicative tool (and that this determines linguistic form), his insights fed into transformational accounts and their successors – extending the theory that prosody can be accommodated in derivational syntactic accounts. In general, this results in extending syntax with focus projections that are contingent on the semantics or discourse; but note that the T-model as presented in 1.2 only allows for a unidirectional flow of derivations. In fact, "externalization (hence a fortiori communication) is an ancillary aspect of language, peripheral to its core nature" (Chomsky, 2015:101). As a consequence, a Chomskyan theory should abide by the notion that meaning of focus depends on a certain syntactic derivation. Though not everyone agrees and many Generativists have adopted Rooth (1992)'s non-syntactic "Alternative Semantics" approach to accommodate focus. Examples of syntactic (deletion) accounts that factor in prosody have been developed by Hartmann (2000) and Winkler (2005), who are in turn indebted to Kuno. Hartmann concludes that a verb can only be deleted if the antecedent is deaccented, while remnants should find a corresponding accented correlate. Note that this resembles Culicover and Jackendoff (2005)'s IL formalisation. Winkler makes a distinction between two different types of ellipsis:

- Sentence-Bound Ellipsis (SBE): The information-structural function of sentence-bound ellipsis is the isolation of contrastive foci or topics. Examples are Gapping, Stripping, PseudoGapping.
- Discourse-Bound Ellipsis (DBE): The information-structural function of discourse-bound ellipsis is to mark the elliptical material as anaphoric or given. Examples are VP-ellipsis, Sluicing, NP-ellipsis.

[after Winkler (2005:37)]

Winkler proposes the "Hybrid Focus Hypothesis of Ellipsis" in which phonological deletion is invoked for SBE while a proform account should be in charge of DBE. Prosodic marking of contrastive focus and topic in SBE is understood as relying on syntactic movement operations (which are by definition sentence-bound). The derivation of a sentence proceeds in two cycles. The first is concerned with the derivation, in which two movement operations take place, which is followed by deletion at the level of PF. The second operation interacts with the first derivation in order to update the appropriate information-structural configuration, assuming a bilateral relation between PF and syntactic formation rules.

Winkler's proposal drops us back into the surface-versus-deep discussion. In the meantime, we have seen that for successful resolution of Gapping-like constructions to happen, discourse representations can be used (see for example Boone, 2014; Culicover & Jackendoff, 2005; Ginzburg & Sag, 2000). Oddly, Winkler implements her ideas in a parallel Jackendoff style (see Figure 1.1) – something she herself acknowledges (see footnote 8 Winkler, 2005:231) – while still taking derivational principles of the Minimalist Program (see Figure 1.2) as her starting point. In her view, LF representations are directly accessible at the level of PF – a position that cannot be accepted if one complies with

Minimalist assumptions. Whatever the validity of her account is with respect to a Minimalist paradigm, it has become clear that a feasible account of ellipsis resolution integrates different – maybe autonomous – levels of representation that are allowed to interact. In that sense, a multi-dimensional sign-based framework such as HPSG may be a more adequate architecture to accommodate multiple levels of representation. Even if a proposed technique does not involve reconstruction of syntactic structure, sensitivity to syntactic parallelism may be included as is argued by Ginzburg and Sag (2000:298). Note though, that the authors are not specific with respect to the relation between focus phenomena (or prosody for that matter) and QUDs, which seems odd since a QUD refers to focused phrases. As noted earlier, Culicover and Jackendoff (2005:539) include a focus constraint in their IL method while further arguing that sufficient syntactic information is available ("matches") during the resolution process.

Recall that we have seen in Chapter 1.3.1 that syntactic constituents do not always coincide with prosodic constituents. Trying to account for this fact and incorporating intonation in a CCG approach, Steedman (1991) makes explicit the integration of prosodic structures with syntactic-semantic structures. The rules of the Combinatory Categorial Grammar are sensitive to intonation for establishing constituency, which would otherwise be rather haphazard. As noted earlier, CCG produces different "trees" for one string. A parser would have to be able to handle such ambiguities all the time. For example, in (35) Steedman shows that different bracketing options are possible in a CCG approach. However, crucially, they all correspond to a distinct intonation contour that in turn corresponds to a certain question that may provide some given information. The question that leaves open the requested phrase, a variable, is referred to by Steedman as "open proposition". Note that this compares with a possible QUD in which the *wh*-phrase represents the variable. A sequence of ") (" marks an intonation and – in CCG terms – constituent break.

(35) a. (They are a good source of) (vitamins).
QUD: What are legumes a good source of?
b. (They are) (a good source of vitamins).
QUD: What are legumes?
c. (They) (are a good source of vitamins).
QUD: What about legumes?

[after Steedman (1991:37)]

The idea is that in order to limit the possible constituent configurations, at least in spoken language, prosodic information should be integrated in the grammar. Steedman proposes that the combination of two syntactic categories via a syntactic "combinatory" rule is added with a restriction that the prosodic categories should also be properly combined. The prosodic categories are subject to prosodic combinatory rules. Similar to the syntactic combination, the assignment of prosody is a matter of functional composition. However, phonological categories are defined as an autonomous (in phonological terms "autosegmental") level of intonational structure. Steedman's proposal offers the possibility to consolidate intonation as structure building rules, pairing phonological and logical form without an intermediate level of representation. Furthermore, speech processing and parsing could be merged into a single process. Steedman uses intonational categories that are based on the autosegmental notation as proposed by Pierrehumbert (1980). A transcription of Dutch intonation (ToDI) has been proposed by Gussenhoven (1988) and is the subject of an ongoing project which can be found on the website todi.let.kun.nl. Converging elements are abstract tonal segments that refer to the relative pitch measured as a function of the fundamental frequency against time. For example, there are two phrasal tones, H and L, denoting high or low 'simple' tones; two boundary tones, written H% and L%, denoting an intonational phrase-final rise or fall; and two target tones written H* and L* denoting a high or low pitch target. Utterances such as (2a) and (2b) repeated here would be transcribed as in (36) (though different pronunciations are possible as reported by Dimitrova, 2012:167).

(36) a. De MAN kocht een boek in LONDEN, en de VROUW in LEIDEN.
H*L LH% LH* H*L%
'The man bought a book in London, and the woman in Leiden.'

b. De MAN kocht een BOEK in LONDEN, en de VROUW een KRANT
H*L H*L LH% LH* LH*
in LEIDEN.
H*L%
'The man bought a book in London, and the woman a newspaper in Leiden.'

Despite the differences in prosodic "contours", the prosodic parallelism in terms of focus is evident. The parallelism reflects intonation contours that are related to the respective QUDs.

(37) a. WIE kocht een boek WAAR?
L*H H*LH%
'Who bought a book where?'

b. WIE kocht een BOEK WAAR?
L*H L*H H*LH%
'Who bought a book where?'

Boone (2014:67) concludes that the fact that the ellipsis clause often seems to answer the QUD is just a reflex of the way focus is assigned. Since he assumes that remnants must be focused independently of the focus structure in the antecedent, he is able to sustain his movement-of-remnants account ("escaping the ellipsis"). Indeed, it appears that neatly aligning focus assignment such as seen between (36) and (37) only holds for structurally parallel cases. The bottom line is that focus assignment is determined by information structure (in the tripartite architecture represented at Conceptual Structure); Boone (2014) would agree with this since he exploits the alternative semantics account by Rooth (1992). What he finds more useful is the notion put forward by Griffiths and Lipták (2014) that contrastive remnants can only be felicitous if their correlate is contrastively focused. This would mean that (38) is not felicitous since *boek* and *Londen* are not contrastively focused – not because the QUD is not properly answered in the right conjunct.

(38) ??De MAN kocht een boek in LONDEN, en de VROUW een KRANT in
H*L LH% LH* LH*
LEIDEN.
H*L%
'The man bought a book in London, and the woman a newspaper in Leiden.'

It is an empirical question to what extent (38) is ungrammatical. To me (being a native speaker), at least, it seems that a felicity condition may be violated. Note that Steedman (1991) would predict (38) to be ungrammatical just as the felicity condition does, only if one (incompatible) QUD is available. Provided that the mechanism is dynamic, though, a prosodic structure may be overruled, giving the opportunity to a secondary QUD – which is in principle possible under Steedman's account. IL as proposed by Culicover and Jackendoff (2005), would set the mechanism to search for an unfocused phrase (here: *kocht een boek*) which can not be properly integrated with the orphan structure. Alternative readings may be guided by the discourse level (c.f. accommodation) – the framework would allow for this – but the authors have no detailed account as to how this would work.

Intonation may help to organise utterances into appropriate constituent configurations, but also into manageable units in order to be perceived and memorised more easily. In cognitive psychology this is known as "chunking" (Crystal, 2010:179) and to get a better understanding of the functional level of description of prosody in ellipsis resolution, we need experimental data. This calls for a theory that can be related to processing, an issue that I will turn to now.

2.6 Mapping to processing

Since the 1970s, theoretical research on Gapping has boomed. Couched in a Chomskyan framework and characterised by derivational levels of representation, the first theories proposed a deletion approach, which is an intuitive choice from a theory-internal perspective. This way, structural parallelism is easily established as the ellipsis site is fully represented, albeit hidden. Therefore, in Minimalism and its predecessors, the operation of deletion and its relative, copying are still applied, with or without some form of phrasal movement. However intricate and admirable in terms of syntactic argumentation, such theories do not have so much to offer a processing account. A representation of a fully-fledged structure may be the ultimate notion that can be used as a starting point for a processing model, but derivational steps cannot easily be mapped to parsing procedures.

Being among the pioneers, Hankamer and Sag (1976) proposed that anaphoric relations cannot be accounted for in mere syntactic terms. Taking the notion of "process" seriously, they further emphasised the importance of discourse. Sag ended up in a constraint-based environment (see for example Pollard & Sag, 1994; Sag & Wasow, 2015) in which the mapping from grammar to procedure is rather straightforward, while different levels of analysis are understood as constituting one dynamic system. An advantage of constraint-based techniques is that we might eventually end up with a theory of ellipsis resolution without needing to resort to the application of different methods for different types of ellipsis.

It seems that theories that assume the lexicon to be the main locus of language-specific grammatical information – proposals of the HPSG, CCG and Simpler Syntax kind – may be the ones that best integrate syntax, semantics, and phonology (including the autosegmental representation of intonation). According to Steedman (1999), such lexicalist models are not only easily paired with computational approaches, but may also be used to show that "there might be a closer relation between the connectionist and symbolist theories than is usually assumed". In other words, such models may be used to relate symbolic representations to lower level neural networks, providing a crucial link to brain activity. Not surprisingly then, such models may be used to better understand the relation between theory and psycholinguistic data using computational methods as a mediating level. Still, CCG proposals have usually been considered to have weak psychological reality. Because of its flexible notion of constituency, a CCG parser would have to deal with (sometimes) numerous possible parse trees of individual sentences. Steedman notes that in principle the CCG approach deals with competence without stipulating how a parser should handle different surface structures (see footnote 4, Steedman, 1996:93). Already in 1991, he stipulates that a parser equipped with instructions related to a functional description of intonation and referential context may be adequate. In other words, intonation contours may guide the interpretation of a certain constituent structure. If one adopted a CCG approach to develop a parser, one should include instructions for handling prosody.

Still, CCG has been used by computational linguists to design semantic parsers; in psycholinguistics they are barely implemented. For example, a computational approach is advanced by Cremers (1993) who argues that the interpretation of coordinate structures – including the ellipsis type Gapping

– is in part "extra-grammatical". In addition to a well-defined description of an expressive grammar formalism, a processing component is included. It is remarkable that in lexicalist accounts on ellipsis such as Cremers (1993), a procedural (computational) implementation may be provided straightforwardly (see for another example Lappin, 1999), while computational accounts embedded in a Minimalist framework need additional stipulations. For example, Kim, Kobele, Runner, and Hale (2011) try to implement the copying account as proposed by Kobele (2015) incorporating "heuristics" that define how a grammar should be applied. This may be contingent on performance factors, and, as a consequence, the authors suggest that heuristics need not be part of grammar *per se*. However, it hardly aligns with the Minimalist T-model which is shielded from interaction with other aspects of cognition. Regardless of the computational theory, typically it deals with meaning representations that may go beyond a structural representation at LF. In general, it may be fruitful for computational research to link theory with practical implementation, possibly opening the door to real-time processing in the brain.

2.7 Summary and conclusions

Theoretical accounts of ellipsis resolution can be broadly summarised by the following quote from Kehler (2000:546) "Syntactic theories recover semantics by copying syntactic material, and semantic theories recover it through a form of anaphora resolution." When it comes to the details of specific approaches, we have seen that each approach works as long as one conforms to specific assumptions of the framework in which the approach is proposed. For example, a movement account (with or without deletion) can only be proposed in one corner of syntactic theory, complying with the presuppositions therein. Consequently, it will be very difficult to link such an account to processing theories since derivational steps such as "first move then delete" have no clear analogue in processing terms. Such theories may be elegant within their framework; at the same time they stand in relative isolation.

Or do they? For example, Boone (2014) can be partly seen as a derivational translation of Indirect Licensing proposed by Culicover and Jackendoff (2005). His reference to the non-hierarchical semantic relation that needs to be in place in order to license Gapping (and thereby Stripping) and Fragments aligns with the idea of semantic parallelism at the level of CS. He then uses syntactic trees to represent such a relation, but in fact he is invoking an independent level of representation – discourse. Within his syntax-oriented framework it is common use to invoke something like D-linking to incorporate such a level. However, the effect is that syntactic theory is augmented with an autonomous dimension that, crucially, may not hinge on syntax *per se*. If one would comply with the model as is depicted in Figure 1.2 on page 6, there is nothing syntactic about discourse – or one should at least have a very lenient view of LF's scope. In other words, Boone presents his account as syntactic but in essence it can be understood as consisting of (at least) two dimensions. The same holds for Winkler (2005) who argues for an interdisciplinary parallel model trying to implement discourse representations within a Minimalist approach. Once again, it is shown that structural approaches try to find ways to discharge Minimalist assumptions.

Finally, the distinction between ellipsis types may only be relevant for theories that involve movement and/or deletion, as it seems redundant for accounts that put less burden on syntactic structure. Theories that involve meaning representations tend to account for Gapping just as they do for other ellipsis types. Given their emphasis on semantic representations, it is relatively easy to link them to computational settings that are typically occupied with meaning representations. Furthermore, accounts that acknowledge autonomous levels are more flexible to incorporate functional levels, which seems essential if one intends to integrate prosody properly.

The simple differentiation between syntactic and semantic accounts has been the driving force behind the current project. The initial idea was to link this differentiation to electrophysiological data. We have only arrived at the end of the second chapter and it already appears that Gapping, and thereby Stripping, cannot be entirely captured in either syntactic or semantic terms as the most promising accounts (will need to) integrate different levels of representations. The next step is to see how these accounts can be connected to existing processing accounts.

CHAPTER 3

Experimental background of Gapping and Stripping

In this chapter I explore the experimental literature on ellipsis. This line of research follows characteristic issues as raised by the theoretical literature to a certain extent. I take into consideration two behaviourally motivated parsing models that are grounded in theoretical insights and that are proposed as a possible link between theory and data. Again, I examine the role of prosody, as well as the relevant ERP components that have been found in relation to the recovery of elliptical structures.

3.1 Objectives and methods of experimental ellipsis research

Successful sentence comprehension requires the listener to parse syntactic, semantic and prosodic information. During parsing, the language processor encodes this information (processes it into memory) and probes it, for example, when earlier processed items need to be retrieved at a later moment. In reading sentences, the same types of information seem to be at play, including prosody – despite the fact that we typically read silently, sound coding appears to be an integral part of the reading process (Pollatsek, Treiman, & Pollatsek, 2015) with punctuation serving to convey prosody to a large extent (Rietveld and Heuven 2001:234; see also the "Implicit Prosody Hypothesis" by Fodor 2002). Assuming that a model for ellipsis resolution should account for the identification, reactivation, and integration of the antecedent, experimental research on ellipsis has entertained the following questions (see also Phillips & Parker, 2014):

- What is the time course of ellipsis resolution?
- To what extent do the antecedent and elided material need to have the same syntactic form?
- Is there syntactic structure at the ellipsis site at all?
- What is the effect of antecedent complexity?
- What is the effect of distance between antecedent and ellipsis site?

These questions may help to show to what extent parsing and interpretation of 'normal' sentences differs from the parsing of elliptical sentences. Different behavioural measures have been used to investigate ellipsis, such as acceptability judgements, comprehension questions, reaction times and reading times.

The questions listed above differ from theoretical questions to the extent that experimental linguists try to develop dynamic models with a focus on the timing of processes. By extension, experimental linguists utilise "judgements" of utterances in a different manner. During acceptability judgement tasks, respondents are for example, asked to specify how acceptable they find certain sentences. Acceptability is usually defined in terms of well-formedness and interpretability of an utterance. In the theoretical literature, argumentation hinges on the grammaticality of sentences. This literature tends to employ a categorical notion of grammaticality: usually, such sentences must be deemed either grammatical or ungrammatical, to sustain a particular theory. By and large, theoretical scholars gain information about grammaticality judgments by consulting their own intuitions. However, it appears that some sentences may have an in-between status, as shown by the examples in (1). Contrasting theoretical linguists, experimental linguists try to explain the gradual difference between these sentences in terms of the relative difficulty to process them which may relate to the time it takes to interpret the utterance.

(1) a. None of the astronomers saw the comet, but John did.
b. ?Seeing the comet was nearly impossible, but John did.
c. ??The comet was nearly impossible to see, but John did.
d. *The comet was nearly unseeable, but John did.

[Arregui, Clifton Jr., Frazier, and Moulton (2006)]

During an acceptability test, speakers are asked for their judgments on a range of sentences. To do so, respondents may be asked to use a rating, or "Likert" scale. Crucially, the way data are collected sets experimental research apart from non-experimental research. During experimentation a variable may be manipulated, for example, the grammaticality of sentence (1a). The manipulated grammaticality can be seen in (1b), (1c) and (1d). It is further required that every other factor that might influence the response is controlled as much as possible. This encompasses a wide variety of factors, ranging from the amount of words in the test sentences to the way participants are instructed, to name just a few. Furthermore, it does not suffice to invite twenty (or so) participants and show them a list of sentences. It may be, for example, important to intermingle experimental sentences with filler sentences if participants are required to remain oblivious to the effect that an experimenter is after. Usually, the sentences are presented in a random order.

Finally, acceptability judgements should be seen as only a first step towards experimental research, as it is not a measure that is generally accepted to be directly informative about underlying processes in the brain. Being an offline task, this method requires participants to think about sentences at a meta-linguistic level – comparable to the theorists approach. This is not the level we are primarily interested in if we want to know how the language system works. Therefore, as a technique it is best-suited for probing intuitions or to pretest stimuli for use in an experiment, that is, to use it as a tool to control variables. For example, for a certain task, experimental sentences should not differ too much in terms of understandability.

Employing a Likert scale in an acceptability judgment task allows the respondents to rate the stimuli on a five- or seven-point scale. Other distributions are possible, but the advantage of an uneven scale is that there is a middle point which may be interpreted as "no opinion" (of course, if a researcher wants to obtain a forced choice an even scale should be used). Further, since participants are likely to avoid the extremes, a seven-point scale may be preferred, if one expects different degrees of acceptability. As an alternative to traditional judgement procedures in the syntactic literature, Bard, Robertson, and Sorace (1996) proposed a Magnitude Estimation (ME) method, which is common in the psychophysics field. In an ME task, participants are asked to estimate the acceptability of a sentence by using their acceptability rating of a different sentence as a unit of measure. However, Sprouse (2011) shows that commutativity does not hold for acceptability judgements: only 20% of the participants were able to decide whether sentences that were equal in terms of difficulty were indeed comparable.

Provided that all stimuli are grammatical, comprehension questions can be used as a method to establish the relative difficulty of utterances. If the researcher is measuring, for example, EEG during reading, it is desired that the participants silently read the presented material. To ensure that participants actually perceive a presented stimulus sentence a comprehension question that targets the content of the stimulus can be presented after the reading task. Although offline tasks – such as acceptability judgements and comprehension questions – have been connected to the relative *processing* difficulty of sentences, in general online measures such as reaction time and reading time are regarded as appropriate behavioural measures to get insight in online structure computation (of whichever information type). Importantly, experimental designs that use online measures (including EEG) usually make reference to a theory (or 'model') that can account for temporal aspects. Before we take a look at psycholinguistic models that have been used to account for ellipsis data, we will briefly return to the discussion of surface versus deep ellipsis.

3.2 Surface versus deep

We have drawn from the theoretical literature that Gapping can be understood as a surface anaphor. This notion has also been advanced in the realm of psycholinguistic approaches. Such approaches have attempted to translate anaphoric relations into "recovery clues". Some research has focused on syntactic parallelism in deep- and surface-ellipsis to put the theoretical proposal of Sag and Hankamer (1984) to the test (e.g. Murphy, 1985a, 1985b; Tanenhaus & Carlson, 1990). A surface anaphor is assumed to bear a linguistic recovery clue by means of the remnant(s) (the sentence subject and often the object). Since a verbal context is required, the possible antecedent is a predicate relation. This relation is assumed to be suitable as a linguistic antecedent. Deep anaphors require recourse to a discourse model.

Murphy (1985b:792) suggests that "the motivation for the deep/surface distinction is related to difficulty in recovering the correct antecedent" and therefore he suggests it is a psychological distinction rather than a linguistic one. He argues that Gapping is part of surface anaphora since a linguistic antecedent, that is literal information about the antecedent, is necessary to interpret the ellipsis. In (2) the relation between subject and object needs to be recovered in order to interpret the clause "and Amy Carl".

(2) a. John duped Bill, and Amy Carl.
b. John believed Bill, and Amy Carl.

[Murphy (1985b)]

In addition to the relation between the antecedent subject and object, "the exact form in which the relation was originally expressed" needs to be retrieved (Murphy, 1985b:803). Hence, the interpretation depends heavily on whether this relation is encoded and remembered as "John duped Bill" or "John believed Bill". Since a structural representation is needed for a successful interpretation, at least for Murphy, this is a powerful reason why Gapping cannot be pragmatically controlled and hence should be regarded as a form of surface anaphor concluding that "conjunction reduction must have a linguistic antecedent by definition, regardless of its recoverability or processing characteristics" (Murphy, 1985b:806).

Note, that – complying with Sag and Hankamer (1984) – for Gapping, an LF representation should be appropriate. Murphy further endorses the notion of copying of syntactic structure, available in short term memory, as a mechanism to resolve surface anaphors. What is confusing is that a statement like "the copying process [...] must have access to the surface form of the antecedent" (Murphy, 1985a:296) may be in accordance with Hankamer and Sag (1976)'s proposal but not with Sag and Hankamer (1984) – given that the surface form of the antecedent is not an LF representation. On the whole, Murphy appears to argue for a parser that may consult syntactic and semantic information types; for him, overt syntactic structure is needed as well as a predicate relation, which might as well be represented in terms of the kind proposed by Dalrymple et al. (1991) and followers.

Using stimuli such as in (3), Tanenhaus and Carlson (1990) showed that, when anaphors are judged to make sense, structural parallelism has a facilitating effect on the speed of processing of both deep and surface ellipsis. This is unexpected if one follows Sag and Hankamer (1984) that would predict that deep anaphors such as "it" in (3c) should not be affected by structural parallelism.

(3) a. Someone had to take out the garbage. (*parallel to condition c and d*)
b. The garbage had to be taken out. (*non-parallel to condition c and d*)
c. But Bill refused to do it. (*deep anaphor*)
d. But Bill refused to. (*surface anaphor*)

[Tanenhaus and Carlson (1990)]

However, Tanenhaus and Carlson found a substantial effect on surface anaphors when no strict parallelism was available, with acceptability ratios being relatively low. In other words, surface anaphors made sense more often in syntactically parallel contexts than in non-parallel contexts. At the same time, parallelism did not affect judgments of deep anaphors. Since an interaction was found between syntactic parallelism and the type of anaphor, they interpreted this finding in favour of Sag and Hankamer. Thereby, Tanenhaus and Carlson sustained the claim that there exists a representational difference between surface and deep anaphors. Instead of a copy mechanism as assumed by Murphy, Tanenhaus and Carlson proposed that antecedents may be linked to an anaphor by means of a "pointer", noting that "copying is not the mechanism by which a surface anaphor would be associated with its antecedent in most current linguistic theories" (Tanenhaus & Carlson, 1990:261). Nonetheless, as we will see in the next section, the debate between copy and pointer approaches would persist.

3.3 Parsing strategies of ellipsis resolution

As mentioned in the first chapter, in the psycholinguistic literature a general distinction can be made between two approaches to sentence processing:

- syntax-first: serial
- constraint-based: parallel

The main motivation for an autonomous syntactic module as proposed by Frazier (for a comprehensive review see Frazier, 1987) was to reduce the burden on working memory load. She countered the assumption of parallel systems, where multiple syntactic analyses for an utterance may be computed and stored at the same time. Furthermore, she suggested that discourse-related factors do not influence an initial syntactic analysis. Her "garden path model" is sustained by abundant evidence that disambiguation of a sentence such as (4) – a famous example – is costly.

(4) The horse raced past the barn fell.

[Bever (1970)]

The syntactic analysis of *raced* is ambiguous since it could be the main verb, or it could be the beginning of a reduced relative clause (where *that was* has been omitted, i.e. *The horse that was raced...*). Parallel models would predict that both analyses compete with each other until the end of the sentence – keeping both sentence structures in memory. A serial model predicts that only one interpretation is being evaluated. By means of the principle known as "Minimal Attachment" (Frazier, 1979) – interpret a sentence in terms of the simplest syntactic structure – *raced* would be treated as main verb. If this interpretation crashed an alternative would be computed in turn. My impression of the serial-parallel discussion, however, is that the division is not clear-cut. Serial models may allow for implementing nonstructural factors during initial stages of syntactic analysis, therefore yielding interactive models. In other words, serial models may exploit different levels of representation in a parallel way. In serial computational approaches, this is most evident. For example, Lewis and Vasishth (2005)'s *Adaptive Character of Thought-Rational* (ACT-R) based serial parser would construct one fully formed analysis of (4) at a time while exploiting different information types. Based on activation decay and retrieval interference, their account utilises strategies that are mainly dependent on probabilities. It may be classified as an intermediate approach, between syntax-first and constraint-based. And, as we will see, it may be related to an ellipsis processing strategy.

A recent representative of a generative parallel constraint-based approach is the computational "surprisal" architecture of Levy (2008) in which "the parser allocates different amounts of resources to different interpretations of the partial input, and difficulty arises when those resources turn out to be inefficiently allocated" (2008:1128). In this model, partial input may be structural or lexical information, that is to say, incoming information is structurally analysed. Levy puts his architecture somewhat in between traditional serial and parallel models since it allows multiple structural variants to facilitate – rather than compete with – the processing of sentences such as (4). The relative difficulty of processing a word is understood as the word's surprisal given its context. In (4), *raced* is highly probable after encountering a noun phrase in English, which amounts to a preferred structure. The author suggests that his surprisal theory should be compatible with reading time accounts based on Lewis and Vasishth's model. However, it is as yet unclear how a surprisal account would handle elliptical structures.

Lewis and Vasishth's model implements a form of predictive parsing that has become known as "left corner parsing": a syntactic structure is built incrementally on a roughly word-by-word basis while little by little predictions may be made about the subsequent structure. Parsing is driven by a bottom-up as well as a top-down (predictive) mechanism. For example, such a parser for Dutch may predict a verb in the second conjunct of (5) as it assumes a structure that parallels the first conjunct.

(5)	De	man	kocht	een	boek	en	de	vrouw	<*e*>	een	cd.
	the	man	buy.3SG.PST	a	book	and	the	woman	<*e*>	a	cd
	'The man bought a book and the woman a cd.'										

Based on behavioural research, it has been suggested that in such sentences the antecedent structure will be activated even before the ellipsis has been detected (Callahan, Shapiro, & Love, 2010). Already at *en* in (5) the processor may use this conjunction as cue to expect a parallel structure upon which *kocht* will be reactivated. Callahan et al.'s evidence for the reactivation of previously processed antecedent information from memory is based on missing verbs and objects. However, it is not clear to what extent other additional phrases such as adjuncts will be reactivated, since they are possible candidates for ellipsis in Gapping and Stripping too.

An important issue concerns the *form* of the reactivated phrase. In accordance with a syntax-first approach, Frazier and Clifton (2001) propose that ellipsis may be resolved by inserting a copy of the missing structure (e.g. *kocht*). This predicts a low processing cost, regardless of the size of the antecedent. This mechanism, known as "Copy α" is assumed to be invoked when elliptical structures are encountered, substituting a step-by-step structure building procedure (for example, the default manner of parsing). Frazier and Clifton's evidence is drawn mainly from reading times that do not show any effect of complexity. Participants performed a self-paced reading task; in this task, they were asked to read text frame-by-frame on a computer screen, pressing a button each time they were ready to move to the consecutive frame. Stimulus texts differed in complexity, as can be seen in (6) (forward slashes indicate the separation of the frames). Crucially, (6a) and (6b) differ in terms of the complexity of the antecedent: that is, in (6b), two clauses need to be retrieved as opposed to one in (6a). Nonetheless, in pairs such as (6), the reading times of the VP-ellipsis "Tina did too" did not differ between (6a) and (6b).

(6) 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.

[Frazier and Clifton (2001, 2000)]

Copy α, however, may not be applicable for all ellipsis types. Notably, Gapping is regarded as an anomaly since it may allow for ambiguous syntactic scope. For example, sometimes the first noun phrase in a right conjunct can be interpreted as a secondary object of the first verb phrase, after which the interpretation crashes. This can be observed in (7), if *de vrouw* is interpreted as object of the verb *zag*.

(7)

De	man	zag	een	boek	en	de	vrouw	<*e*>	een	cd.
the	man	see.3SG.PST	the	book	and	the	woman	<*e*>	a	cd

'The man saw a book and the woman a cd.'

Gapping thus prevents α from being straightforwardly determined and as a consequence, Gapping should be subject to a different parsing approach. This reminds us again of the surface versus deep discussion. Two different parsing strategies could be assumed for the deep-ellipsis in (8a) on the one hand and the surface ellipsis in (8b) on the other.

(8) a. Sarah bought a book. Tina did too. (VP-ellipsis)
b. Sarah bought a book. Tina too. (Stripping)

Just as Gapping, Stripping seems to allow for ambiguous readings changing the syntactic scope. "Tina" in (9b) may be agent or patient. Therefore, a copy mechanism would not be applicable to Stripping.

(9) a. Sarah broke up with her boyfriend last May. Tina did too. (VP-ellipsis)
b. Sarah broke up with her boyfriend last May. Tina too. (Stripping)

More recently Clifton. Jr. and Frazier (2010) have proposed that ellipsis (in general) may be constrained by structural and discourse conditions. Structural conditions would be provided by the grammar while the application of discourse conditions are assumed to be a quality of the processor. Maintaining a copy mechanism (in subsequent work this notion evolved into structure "sharing" (Frazier & Clifton, 2005) and "recycling" (Arregui et al., 2006;

Frazier, 2013) of syntactic structure, would mean that its applicability is determined by the grammar. At the same time, it is once again clear that a characterisation of ellipsis requires an interactive approach as the choice of the antecedent in (9b) requires access to different information types. Something to keep in mind in relation to copying or sharing structure is that memory for surface aspects (e.g. syntactic structure) appears to be relatively short-lived, in contrast to semantic information which can still be accessed after longer stretches of discourse. Garnham and Oakhill (1987:614) refer to this phenomenon as "one of the best-established results in the psycholinguistic literature.". This compares with the short-lived nature of the phonological code for a word, which becomes available when the word is accessed (Levelt, 1999) but decays within four to seven words (Baddeley, 2012; Tanenhaus, Carlson, & Seidenberg, 1985).

Others have argued that, in general, accessing a copy of some structure should take more time since a serial search must be undertaken. For example, Martin and McElree (2008) have proposed an alternative view of ellipsis resolution that involves step-by-step structure building, while a more fully interpreted discourse representation of the antecedent is accessed and integrated. Implemented as a memory-based content-addressable pointer mechanism, it would involve direct access of the antecedent, suggesting that the speed of interpreting the ellipsis does not depend on antecedent complexity. Taking Copy α as competing account, Martin and McElree (2008) propose that a pointer should not be limited to mere syntactic structure; rather it might as well be linked to a "fully interpreted discourse representation" (2008:883). Their approach not only bears resemblance to pro-form theories of ellipsis, but it is also closely connected to the ACT-R approach as proposed by Lewis and Vasishth (2005). Notably, Callahan et al. have suggested that some aspects of predictive parsing may be underpinned by such a cue-based method.

In one experiment, Martin and McElree asked participants to determine the acceptability of sentences as in (10) and (11) (choices: yes/no). Again, forward slashes signal frames of phrase presentation, but in this experiment the frames were presented at predefined moments (i.e. it was not a self-paced task). Participants were required to respond after every sentence.

(10) *Simple antecedent*

a. The history professor / understood Roman mythology, / but the principal / was displeased to learn that / the over-worked students / attending summer session / did not.

b. *The history professor / understood Roman mythology, / but the principal / was displeased to learn that / the overly worn books / used in summer session did not.

(11) *Complex antecedent*

a. The history professor / understood Rome's swift and brutal / destruction of Carthage, / but the principal knew the over-worked students / attending summer session / did not.
b. *The history professor / understood Rome's swift and brutal / destruction of Carthage, / but the principal knew the overly worn books / used in summer session / did not.

[Martin and McElree (2008)]

Martin and McElree found that participants' accuracy on the acceptability task and response time in condition (11) was comparable to condition (10). The authors hypothesised the response time to be a reflection of the time needed to retrieve an antecedent and to interpret it at the ellipsis site. Since the response time was not affected by the complexity of the antecedent, they used this finding as evidence for a pointer mechanism during which a pointer directly accesses an antecedent. In their comprehensive review of experimental approaches to ellipsis processing, Phillips and Parker (2014) note that to establish the acceptability of these test sentences, it suffices to link the subject of the embedded sentence ("the over-worked students" versus "the overly-worn books") to the head of the antecedent VP "understood" while the object content may be disregarded. They further question the statistcial power of (Frazier & Clifton, 2001)'s Copy α study, a problem that has been overcome in follow-up experiments reported in Frazier and Clifton (2005).

Additional support for Martin and McElree's account is based on non-effects of distance between antecedent and ellipsis. Tested sentences as in (12) showed that distant antecedents yielded poorer accuracy of acceptability judgement, but crucially, the processing speed was not compromised.

(12) *Near antecedent*

a. The editor / admired the author's writing, / but the critics / did not.
b. *The editor / admired the author's writing, / but the binding / did not.

(13) *Distant antecedent*

a. The editor / admired the author's writing, / but everyone / at the publishing house / was shocked to hear that / the critics / did not.
b. *The editor / admired the author's writing, / but everyone / at the publishing house / was shocked to hear that / the binding / did not.

[Martin and McElree (2008)]

What strikes me most is that both Copy α and the cue-based mechanism predict comparable results: no processing cost. At least for the cue-based method, this prediction holds if no additional referents intervene between antecedent and ellipsis. "Interpreting additional material decreases the specificity of retrieval cues" which may amount to "*cue overload* where the cues that make contact with the correct constituent in memory are insufficient for successful retrieval" (Martin & McElree, 2011:330). All in all, this leaves us with inconclusive evidence for structural complexity effects in particular. In the present study, we will explore the possibility of differentiating between the proposed mechanisms by means of electrophysiological data.

3.4 Prosody

As we have seen in the discussion on Copy α, Gapping-like constructions may lead to ambiguous readings. A seminal behavioural study on the effects of parallelism on the interpretation of Gapping is Carlson (2001). Maybe unsurprisingly, an ambiguous reading is dependent on (*i*) the thematic role restrictions of the elided verb, that is, the possible object the elided verb selects; (*ii*) the prosodic contour; and (*iii*) the discourse context if available. Using a written questionnaire using sentences as in (14), Carlson found that participants favour a Gapping analysis when a parallel structure between arguments is available (encouraged by italicising them). The percentage of Gapping interpretations is indicated in brackets.

(14)
a. *Alice* bakes cakes for tourists and *Caroline* for her family. (81%)
b. Alice bakes *cakes* for tourists and *brownies* for her family. (3%)
c. *Josh* visited the office during the vacation and *Sarah* during the week. (40%)
d. Josh visited *Marjorie* during the vacation and *Sarah* during the week. (4%)
e. *Dan* amazed the judges with his talent and *James* with his musicality. (21%)

[Carlson (2001)]

Clearly, italicising parallel phrases influence interpretation, but note that still a preferred reading of conjoined objects is apparent in condition *c*. In other words, prosody as imposed on written input is an additional factor. However, it cannot overrule an alternative thematic fit imposed by the verb.

In a follow-up experiment, Carlson tested the hypothesis of Minimal Attachment as discussed in the preceding section (on page 54). Based on this principle she derived the hypothesis that the simplest legitimate syntactic analysis of an input is preferred during parsing. She compared this to a parallelism constraint that dictates that the most parallel analysis of a conjoined structure is preferred, which holds that the parser should look for similar syntactic roles. Carlson used stimuli such as explained in (15) and presented them auditorily. Immediately after presentation of a sentence, a question appeared on-screen asking the participants to choose the best paraphrase of the sentence they had heard.

(15) a. Josh visited the office during the vacation and Sarah during the week. (*baseline prosody*)
b. [BOB insulted the guests during DINNER] and [SAM during the DANCE]. (*cooperating Gapping prosody*)
c. Bob insulted [the GUESTS during DINNER] and [SAM during the DANCE]. (*cooperating Non-Gapping prosody*)

[Carlson (2001)]

A baseline prosody predicts that when the ambiguous DP "Sarah" contrasts equally with either the first conjunct argument ("Josh") or "the office", the prosody will be compatible with both the Gapping and Non-Gapping interpretations of the sentence. Contrastingly, cooperating prosody can bias the interpretation toward the Gapping or Non-Gapping analysis.

The Gapping response rate for condition *a* was 38% and is comparable to the 40% seen in the written study. Carlson concludes that a special Gapping prosody is not necessary for choosing a Gapping interpretation. The status of baseline was further corroborated by the Gapping response ratios for conditions *b* (44%) and *c* (28%), putting the baseline (almost) in-between. Again, a paralleled focused argument could not overrule a minimal attachment approach of the parser.

3.5 Electrophysiological research on ellipsis

As I have explained in Chapter 1.3.3 event related potentials (ERPs) can be measured to investigate the interplay between semantic, syntactic and (to a lesser extent) prosodic processes. During an ERP experiment, the EEG signal is analysed relative to specific time points in the experimental presentation. For example, when the time point at which a stimulus appears on-screen, researchers investigate how the brain activity responds to that particular stimulus. The average signal per condition per participant per electrode is averaged to get a "Grand Average" per condition (and electrode). The difference between the Grand Averages per condition is what is called event-related potential (ERP). Typically, in reading experiments, sentences are presented word-by-word. The onset of the presentation of a critical word is then taken as measure point for the onset of the ERP. Five main markers have been established in the literature: CPS, ELAN, LAN, N400 and P600 (see Table 1.1).

Given that the time resolution of electrophysiological methods is very precise, the focus in ERP research on ellipsis is on the time course of processing. Three steps in this time course may be considered:

- identification of missing structure
- reactivation of the antecedent
- integration of the retrieved antecedent

These steps are derived from two ERP studies on Gapping. The first study, carried out by Kaan, Wijnen, and Swaab (2004), investigated the time course of identification and resolution of verb gaps in English. Using sentences as in (16) they manipulated the plausibility of critical noun phrases ("the hammer") that followed a verb gap. For convenience, the elided verb is indicated between <>.

(16) a. Ron took the planks, and Bill <took> the hammer.
b. Ron sanded the planks, and Bill <sanded> the hammer.

[Kaan et al. (2004)]

Sentences were presented word-by-word and the task of the participants was to click either 'GOOD" or 'BAD" in order to rate the stimuli as a means of acceptability task.

The authors hypothesised that the determiner of the critical NP would be the first possible point for the processor to detect a missing structure – a verb. Since they recorded a centro-posterior negativity they suggested this to be a variety of an ELAN effect which may be connected to phrase structure violations. In addition to the early negativity, a positive fronto-central deflection between 300-500 ms was measured, which they cautiously ascribe to the retrieval of the preceding verb information.

What is problematic is that Kaan et al. compared the critical determiner to any determiner appearing in any sentence in this study. This leads to the appearance of unexpected effects, since the determiners that were used as baseline appeared in different positions in a sentence. As such, the established negativity does not make a strong argument for the detection of missing structure. Despite this, the authors propose a general mechanism of syntactic persistence reminiscent of the Copy α routine that extends beyond ellipsis: "the grammatical structure of a sentence can be stored in the working memory as an autonomous entity, and re-accessed in subsequent processing" (2004:590).

Kaan et al. reasoned further that if a missing verb were reactivated (they also refer to this as "reconstruction" of missing structure, here), integration difficulty would be apparent at the critical noun. With respect to this difficulty an N400 was expected, and was found. This was interpreted as evidence for reconstruction of the antecedent at "hammer" or just after detection of the gap (at the determiner). Note that an N400 cannot be interpreted in terms of the syntactic form of the ellipsis construction, but in terms of the lexical representation of the antecedent and possibly the relative difficulty of integrating this antecedent. In addition to their expected N400 effect, a P600 was found. They suggested that, possibly, the semantic anomaly (e.g. "sanded the hammer") may have induced a process of syntactic revision. This seems odd since there is no difference in terms of syntax between plausible and implausible reconstructed phrases, that is, the anomalous verbs do not yield a different syntactic construction. Since a P600 may be larger when violations are task-relevant (see for a discussion Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014), it may be triggered by an acceptability task (as was used by Kaan et al.). Possibly due to meta-linguistic processes that are at work during an aceptability task, the authors can not be sure that the P600 was a reflection of a combined syntactic and semantic integration difficulty.

Kaan, Overfelt, Tromp, and Wijnen (2013) is the only ERP study in which Dutch sentences containing Gapping are compared with "No Gapping" counterparts. Just as the earlier study on English described above, only a verb was elided. They focused on the moment when the ellipsis site in Gapping constructions is detected and the moment when the antecedent is accessed and integrated. However, in contrast to the English study, they investigated to what extent an ellipsis site can be anticipated, rather than focusing on the question of whether there is syntactic structure at the ellipsis site. The move towards anticipation rather than trying to establish some kind of structure coincides with the current fashion of probabilistic methods (as put forward by for example Hale, 2011; Levy, 2008). But also, it appears to be very difficult to demonstrate that there is structure at an ellipsis site, despite the wide range of experimental techniques that have been used in attempting to do so (see for an overview and discussion Phillips & Parker, 2014). In Kaan et al. (2013), the left anterior negativity found in Kaan's earlier study is reinterpreted in terms of expectations. If the parser does not anticipate an elided verb, this may result in processing difficulty. However, in the case that ellipsis is expected, no such ERP component should be found. An example set of the stimuli used in this experiment is shown in (17). Stimuli were presented on-screen word by word and participants were again asked to determine after each sentence whether it was good or bad. In (17), we see four conditions of which (17a) and (17c) are of the most interest: these allow us to compare plausible Gapping with plausible No Gapping sentences while the critical measure point is the NP "de bloemen". However, the status of the second clauses differs. In condition *a*, we see two conjoined main clauses, whereas in condition *c*, a main clause is followed by a subordinate clause.

(17) a. Anouk zond de kaart aan haar vader, en Julia <zond> de
Anouk sent the card to her father, and Julia <sent> the
bloemen aan haar moeder.
flowers to her mother.
'Anouk sent the card to her father, and Julia the flowers to her mother.' (*Plausible Gapping*)

b. Anouk schreef de kaart aan haar vader, en Julia <schreef> de
Anouk wrote the card to her father, and Julia <wrote> the
bloemen aan haar moeder
flowers to her mother.
'Anouk wrote the card to her father, and Julia the flowers to her mother.' (*Implausible Gapping*)

c. Anouk zond de kaart aan haar vader, terwijl Julia de bloemen
Anouk sent the card to her father, while Julia the flowers
aan haar moeder stuurde.
to her mother shipped.
'Anouk sent the card to her father, while Julia shipped the flowers to her mother.' (*Plausible control for a*)

d. Anouk schreef de kaart aan haar vader, terwijl Julia de
Anouk wrote the card to her father, while Julia the
bloemen aan haar moeder stuurde.
flowers to her mother shipped.
'Anouk wrote the card to her father, while Julia shipped the flowers to her mother.' (*Control for b*)

[Kaan et al. (2013)]

The authors argue that in a subordinate clause (as in condition *c*) verb Gapping is prohibited, and crucially, no verb Gapping is expected after "Julia" since a reader (or listener) may predict the location and appearance of a verb as soon as the conjunction has been processed. Therefore, "terwijl" indicates that the verb will be appearing at the end of the clause, while "en" introduces a main clause requiring a verb right after the first phrase. According to Kaan et al., conditions *c* and *d* therefore are proper control conditions as no verb is expected before the NP "de bloemen". The plausible versus implausible conditions were added to investigate when the elided verb is semantically integrated – similar to the study in English discussed above.

The critical measure points were the determiner and the noun in the phrase *de bloemen* 'the flowers'. To analyse effects on the determiner the grand averages of conditions *a* and *b* were collapsed and compared with collapsed conditions *c* and *d*. No effects could be found. However, only for participants who performed poorly on the task, i.e. participants who had problems determining that a sentence such as (17b) was bad, ERP effects at the determiner showed a negativity between 400-600 ms after onset. The authors suggest that this negativity is a later instance of the ELAN found in their earlier study. They regard it as a LAN that has been suggested as being an index of prediction strength, as is proposed by Lau, Stroud, Plesch, and Phillips (2006).

Similar to the English study, a P600 effect was found in the implausible Gapping condition at the noun. Notably, the P600 was also apparent in plausible Gapping constructions. However, the N400 was only slightly visible on the Cz electrode and there was no significant effect here. Concluding that the integration of the elided verb is a relatively late process (i.e. 600 ms after presentation of the noun), the authors suggest that the involvement of a resolution mechanism for Gapping is similar to integrating a *wh*-phrase object with its verb as in, for example, "Which book did you buy?"(Kaan, Harris, Gibson, & Holcomb, 2000). Upon encountering the verb "buy", the earlier processed object "which book" can be integrated. Note though that this example differs qualitatively from ellipsis processing, since a *wh*-phrase object always awaits obligatory integration, in contrast to antecedents of ellipsis. Kaan et al.'s findings seem compatible with the integration part of Brouwer and Hoeks (2013)'s "Retrieval-Integration" account that I introduced in Chapter 1.3.3. It is, however, still unclear what the form of the retrieved antecedent might be.

Kaan et al. assumed that an absence of a LAN would be evidence of a top-down approach during which the parser already reactivates antecedent information at the connective *en* ("and") since it expects an ellipsis. This would contrast with a bottom-up approach which would infer an ellipsis site upon encountering missing structure. A LAN would then be a sign of gap detection. Some readers may take issue in general with the principle of "absence of evidence is evidence of absence". In other words, absence of a LAN may well mean absence of a bottom-up approach but this should be taken to entail the existence of a top-down approach. Furthermore, it is striking that the authors overlook the fact that Gapping is an optional process. That is, there could be a verb following *Julia*, one that contrasts with the verb in the first conjunct. What is expected then by the parser, is verbal information, but not verb Gapping exclusively. In addition, the object can be elided as well, leaving the parser uncertain as to how much information should be reactivated.

As shown in (17), a comma was used to make explicit a separation between the two clauses. Although participants were encouraged to interpret a sentence such as (17a) as a parallel coordination of two clauses, it might be that some participants applied a process reflecting minimal attachment yielding an object coordination of "aan haar vader en Julia". Note that, although such a parsing strategy was certainly not possible in all experimental stimuli, it could have affected the results. For example, we could interpret the observed LAN as a reflection of the parser resetting the "Minimal Attachment" principle. By and large, the LAN cannot be straightforwardly linked to any expectations the parser might have, let alone for plausible, say, grammatical instances of, Gapping.

In her doctoral thesis, Dimitrova (2012) reports two auditory ERP experiments on Gapping using prosody as modulation to disambiguate structurally ambiguous constructions. Using sentences as in (18) she measured ERPs at the subject of the second conjunct. Sentences were presented in isolation in one experiment and presented with a preceding question (that biased towards the applicable prosody of the following sentence) in another. "Good thematic fit" means that a minimal attachment applies, that is, a conjunction of two clauses is dispreferred.

(18) a. John invited PETER on Monday and MARTIN on Tuesday. (*preferred-good thematic fit-no Gapping prosody*)
Discourse: When did John invite the boys?
b. JOHN invited Peter on Monday and MARTIN on Tuesday. (*dispreferred-good thematic fit-Gapping prosody*)
Discourse: When did the boys invite Peter?
c. JOHN peeled the orange with a knife and MARTIN with his hands. (*enforced-poor thematic fit-Gapping prosody*)
Discourse: How did the farmers peel the orange?
d. John peeled the ORANGE with his knife and MARTIN with his hands. (*anomalous-poor thematic fit-no Gapping prosody*)
Discourse: What did John peel?

[Dimitrova (2012)]

In the trials where no preceding question was present to provide disambiguating discourse context, Dimitrova found in condition (18b) a marginal right-lateralised negativity (400-700 ms), followed by a positivity (700-1,000 ms) relative to (18a). In sentences with poor thematic fit (18c)-(18d), Gapping prosody elicited a broadly distributed negativity (400-700 ms). Nongapping prosody in (18d) triggered an anterior-central negativity (700-1,000) and a posterior positivity (700-1,300 ms) as compared to (18c).

When discourse contextualising questions did precede the sentences, a centro-posterior negativity (400-700 ms) was elicited irrespective of thematic fit in Gapping readings ((18b) and (18c)). In good thematic fit sentences, the Gapping bias as established by the preceding question yielded an anterior positivity and posterior negativity in two subsequent time windows: 700-1,000 and 1,000-1,300 ms.

Dimitrova suggests that the interpretation of an accented ambiguous element such as "MARTIN" in a Gapping reading yields an N400-like component – apparent in both experiments. She considers this component "to be related to the activation of verb phrase information and the assignment of a subject role (rather than an object role) to the accented element" (2012:228). She found the N400 in sentences with and without a biasing context. Just as we have seen in Carlson's study (2001), a Gapping reading for a sentence like (18b) is dispreferred. Dimitrova relates the established P600 to the reconstruction of a more complex (and dispreferred) structure, yielding two conjoined clauses.

However interesting these results are, it seems odd to attribute the negativity found in (18c), as compared to (18d), to the same underlying procedure as apparent in condition (18b). "MARTIN" in (18c) is compared with an anomalous condition. In other words, we cannot be entirely sure that the effect found was due to Gapping prosody alone; rather, they could be interpreted in the opposite direction – as is usual in experiments using anomalous conditions. To my knowledge, there is only one additional ERP (reading) study on ellipsis – although not on Gapping. In favour of the cue-based mechanism, the study of Martin, Nieuwland, and Carreiras (2012) shows that retrieval interference (possibly due to "cue overload") is reflected by a negativity between 400 and 1,000 ms after stimulus onset. Note, though, that this study concerns Noun Phrase ellipsis. This is crucially different from Gapping since no predicate relation is involved. Spanish sentences such as in (19) were presented to participants word by word. Some sentences (60% of the trials) were followed by a comprehension question. Between brackets, gender of a noun is indicated. The NP-ellipsis "another" has to match with the correct gender of the antecedent. In all sentences, there is an intervening noun (an attractor) that is structurally unavailable as antecedent.

(19) a. Marta se compró la camiseta que estaba
Marta REFL buy.3SG.PST DET.F t-shirt REL be.3SG.PST
al lado de la falda y Miren cogió
PREP+DET.M next PREP DET.F skirt and Miren take.3SG.PST
otra para salir de fiesta.
another.F to go.3INF PREP party.
'Marta bought the t-shirt (fem.) that was next to the skirt (fem.) and Miren took another (fem.) to go to the party.' (*correct attractor-same*)

b. Marta se compró la camiseta que estaba
Marta REFL buy.3SG.PST DET.F t-shirt REL be.3SG.PST
al lado del vestido y Miren cogió otra
PREP+DET.M next PREP DET.F skirt.F and Miren take.3SG.PST
para salir de fiesta.
another.F to go.3INF PREP party.
'Marta bought the t-shirt (fem.) that was next to the dress (masc.) and Miren took another (fem.) to go to the party.' (*correct attractor-different*)

c. *Marta se compró la camiseta que estaba
Marta REFL buy.3SG.PST DET.F t-shirt REL be.3SG.PST
al lado de la falda y Miren cogió
PREP+DET.M next PREP DET.F skirt and Miren take.3SG.PST
otro para salir de fiesta.
another.M to go.3INF PREP party.
'Marta bought the t-shirt (fem.) that was next to the skirt (fem.) and Miren took another (masc.) to go to the party.' (*incorrect attractor-same*)

d. *Marta se compró la camiseta que estaba
Marta REFL buy.3SG.PST DET.F t-shirt REL be.3SG.PST
al lado del vestido y Miren cogió otro
PREP+DET.M next PREP DET.F skirt.F and Miren take.3SG.PST
para salir de fiesta.
another.M to go.3INF PREP party.
'Marta bought the t-shirt (fem.) that was next to the dress (masc.) and Miren took another (masc.) to go to the party.' (*incorrect attractor-different*)

The increased negativity found in (19b) at the NP-ellipsis is considered as a retrieval interference. In the ungrammatical sentences (19c) and (19d), a sustained negativity was found. The authors concluded that "structurally unavailable noun phrases are at least temporarily considered for grammatically correct ellipsis" (2012:1859). This would mean that cues are stored and activated regardless of the syntactic structure they originated from.

3.6 Summary and conclusions

In this chapter, I have highlighted the results of previous experiments related to structural complexity and prosody with a focus on Gapping. The experimental literature on ellipsis very much reflects the characteristic issues raised by the theoretical literature. Grounded on theoretical insights, models have been proposed as a link between theory and data. Two of these, Copy α and the cue-based mechanism, reflect to some extent the divide between syntax-first and constraint-based approaches. Although it is tempting to use these models to estimate the extent of syntactic structure available at the ellipsis site, I would like to quote Phillips and Parker (2014:15)'s conclusion that "caution is required in mapping findings about timing of ellipsis resolution onto theories of the representation of ellipsis constructions" .

Carlson (2001)'s experiments emphasise the role of prosody during Gapping resolution. Although apparent, prosody does not outweigh the influence of verbal information. It would therefore be interesting to see how different prosodic contours modulate conjunctions that are not ambiguous between Gapping and Non-Gapping. Just as in Dimitrova (2012), an ERP experiment could be designed to do so. As an alternative explanation of differences imposed by different prosodic contours, a lack of parallel intonation could be interpreted by the parser as a cue that the unaccented argument may be considered as possible structure to elide. Prosody, then, would help make predictions about upcoming structure.

Kaan et al. (2004)'s finding of ELAN in combination with a positivity between 300-500 ms at the determiner may be a reflection of a retrieval process, but we cannot be sure about the form of the retrieved material in this study. Although the authors suggest that the antecedent is reconstructed at that point as part of a general mechanism of syntactic persistence, retrieved information might be of another information type which is integrated once the object is processed. This integration process was tested more adequately in their follow-up study, comparing Gapping and similar Non-Gapping constructions. The P600 found in that study could well reflect an integration process. However, it remains unclear on which information type(s) of the antecedent this process was operating. For example, instead of dealing with fully-fledged structure, the integration phase may be confronted with a more fully-interpreted chunk.

It might be difficult to ascertain either information type in the ellipsis site. However, if we can tease apart the different predictions of the models discussed above, we might end up with a indication of how to map neuronal activity to representations proposed by the theoretical literature. In the next chapter, I will argue that this might be possible.

CHAPTER 4

Setting the stage

In this chapter I discuss the mapping between existing theoretical insights and actual processing. I arrive at a comparison of Copy α and the cue-based mechanism with respect to the timing of processes of retrieval and integration. Since individual differences may lead to differences in (amplitudes of) ERP components and may be ascribed to natural variability in the capacity of human working memory, I further propose a suitable working memory test. I conclude with hypotheses and possible results.

4.1 Bridging theoretical and experimental research

4.1.1 Introduction

Linguistic research at the Leiden University Centre for Linguistics, the place where the current study was carried out, may be broadly described as the study of structure and variation among the world's languages. At this institute three types of linguists ranging from theoretical and descriptive to experimental can be found in different workspaces: the so-called armchair, the field, and the lab. All working towards an understanding of human language, it seems at times that their insights are difficult to reconcile. This chapter[1] describes a framework for those linguists who view language ultimately as a cognitive system.

The following analogy may be used to show that the division of workspaces need not lead to a segregation into distinct linguistic fields per se. Imagine Anne, sitting in her garden chair, noticing that the ants in her garden are walking faster as the temperature increases. While sitting there, she comes up with a function rule for this phenomenon. Her neighbour Eddy embarks on a jungle trek in South America and tries to apply the function to the Amazonian ants, without any success. However, he does notice that there is a high level of humidity. He decides to build a database in which he lists facts about temperature, humidity and walking speed of the different ants he finds. Already questioning the domain and range of Anne's function,[2] their mutual friend Onno checks the limits of Anne's proposed function in his beloved botanical garden, manipulating both temperature and humidity as possible factors. He further relates his findings to the physical properties of several kinds of ants. Finally, the three friends arrive at an integrated theory of the ant's walking speed.

While the link between abstraction and observation in the analogy is pretty straightforward, it is clear that three methods of investigation have all contributed to the understanding of the ant's behaviour. At the same time, predictions stipulated by their *shared* theory can easily be tested in different research domains. The idea is that the complementary aspect benefits empirical research, which encourages linguists in the armchair, field and lab to better understand each other, given that their research goal is the same. In what follows, I will show that different methods of data collection and different levels of analysis need not divide linguistics into separate fields.

[1] Section 4.1 has been published in Reckman, Cheng, Hijzelendoorn, and Sybesma (2017) and are slightly edited for this dissertation.

[2] This function appeared in an actual math assignment in the 1980s in the method *Getal en Ruimte* (Noordhof Uitgevers). To the amusement of the class, an ant could end up walking backwards at some degree below zero.

4.1.2 What is at stake?

Within a generative approach, the grammar system is usually assumed to be a static entity of knowledge that resides in the brain and that interacts with a processing system containing comprehension and production mechanisms. Ever since Chomsky's seminal work *Aspects of the Theory of Syntax* (Chomsky, 1965), the division between competence (grammar system) and performance (processing system) has fuelled linguistic analysis. However, 50 years later it also appears that the hypothesised division has constituted an obstacle for linguists who aim to link linguistic theory to neuro- and psycho-metrical data. This chapter will consider two apparent issues. Firstly, unlike the properties of physical phenomena such as the walking speed of an ant, we (still) lack a device that can objectively and directly measure the properties of a cognitive phenomenon such as "grammar". Although in the past decades a division of data collection has usually been linked with two separate language systems, we will see in section 4.1.3 that while such a division may be ideal in terms of theory, it is obscure in practice. The second problem concerns the linking of two separate systems. The question here is, how should the interaction between a grammar and separate processing system be defined? A possible (beginning of a) solution of this problem will be the topic of section 4.1.4.

4.1.3 Methods of measurement

Offline versus online data collection

Traditionally, linguistic theory finds its basis in categorical distinctions conceived of and assessed by means of introspective judgements – either those provided by the linguist, or informally collected by asking colleagues at work, conferences or other meetings. The use of judgement data that is collected by means of controlled experiments (either through web-based tools, in the field or in the lab) has been embraced by some, but it is still frowned upon by others (see for a discussion the special issue of *Theoretical Linguistics, 33* (3), 2007). A hypothesised split between introspection and experimentally collected judgements seems, however, untenable: although it has usually been taken as fact that introspection is a reflection of linguistic competence, operating beyond any kind of performance, one could say that this method of research forms one end of an empirical continuum, ranging from well-informed individual offline judgements gathered from colleagues, to online measures of a naive group of people taken in highly controlled experiments. In addition, corpus data collected in the field (i.e. systematic collections of naturally occurring texts of both spoken and written language) may be used to quantify linguistic phenomena, which may be useful at any level of the empirical continuum. For example, an experimenter may need to extract data from a corpus to check for possible confounds in a stimulus set. Whichever method is used, the commonality between linguistic researchers is that they try to generalise their results to the speech community.

During an offline judgement task, a participant (or a single linguist) responds to a certain linguistic stimulus with no time restrictions. Meanwhile, online measures such as reaction time, eye movements, brain potentials – to name a few – may give insight into online language structure computation. While a considerable part of linguistic theory is unconcerned with online processes, it is desirable that, if a theory purports to have computational strength, it is at least able to specify how computations are implemented during online language use. In this sense, experiments that measure reaction time, for example, may add to a (computational) theory of language just as judgement data do. However, it is unwise to use reaction time data to compare two theoretical constructs that both lack hypotheses about timing to start with.

On the assumption that offline responses reflect the representations of the grammar and online responses reflect processing mechanisms, Lewis and Phillips (2015) note that frequent misalignments between offline and online responses should be apparent. By "alignment" they mean the extent to which constraints of language processing are the same as those imposed by the grammar. Take, for instance, a garden-path sentence such as (1).

(1) Colleagues sent the invitation to Crit's retirement party were happy.

Only if the reader or listener is given enough time, an initial computation, in which *colleagues* is subject of *sent*, can be revised. Online the sentence may be judged ungrammatical, contrary to an offline response.

While it is indeed the case that misalignments exist, the authors effectively show that specific types of misalignment between online and offline responses amount to specific stages of computation. Crucially, the misalignments seem predictable. The authors therefore claim that:

> [...] online and offline representations are the product of a single structure-building system (the grammar) that is embedded in a general cognitive architecture, and misalignments between online ("fast") and offline ("slow") responses reflect the ways in which linguistic computations can fail to reflect the ideal performance of that system. (Lewis & Phillips, 2015:39)

Treating different methods of data collection as lying on a continuum naturally corresponds to Lewis and Phillips's view that representations of one language system can be investigated using different measures. As different methods target the same representations, the object of study is one system. The advantage of apprehending a single system is twofold. Firstly, it opens the door to incorporating gradient patterns that have been reported in both theoretical (usually indicated by question marks) and experimental (yielded by measurement type) research. Secondly, no separate account is necessary as to how those representations are identified during comprehension and how they are assembled during production. Future research is needed to confirm that a single system can carry out both comprehension and production tasks. In the remainder of this chapter I will explore which levels of description we need in order to define a cognitively-motivated language system such as this.

Online data in ellipsis research

Elliptical sentences, such as the Gapping example in (2), are interpreted by a process of "retrieving" and "integrating" earlier mentioned information (here: *bought a book* denoted by <*e*>).

(2) Eva bought a book in the shop, and Agnes <*e*> in the supermarket.

Interestingly, we can arrive at an interpretation of this sentence within a linguistic context without immediately available linguistic form. Though ellipsis is a multidimensional phenomenon since syntactic, semantic and prosodic constraints apply, theoretical approaches usually take one of these dimensions as a starting point to account for the nature and "recoverability" of the antecedent. For example, syntactic accounts generally represent the elided content as a fully-fledged structure at some point during the derivation, while semantic accounts would recognise the ellipsis as a more fully interpreted representation. A further issue concerns "licensing" which relates to the question of when ellipsis is allowed: which elliptical structures are well-formed? Although hybrid theories exist, an integrative theoretical account which incorporates syntactic, semantic and prosodic constraints of even one type of ellipsis is still to be developed, let alone a unified account of the phenomenon as a whole. Notably, Cremers (1993) argued that the interpretation of coordinate structures – including the ellipsis type as seen in (2) – is in part "extragrammatical", linking to a processing component.

In the psycholinguistic literature on ellipsis comprehension, it has been suggested that acceptability of ellipsis may depend on the amount of repair that is required to resolve omitted structure that does not exactly match the antecedent structure (Arregui et al., 2006). For example, using sentences such as (3) (repeated from Chapter 3.1) an acceptability decline ("gradience" if you will) can be observed, (3a) being judged most acceptable, and (3d) least acceptable. The example shows that a decline correlates with the relative difficulty the processor experiences in recovering the phrase *see the comet* in the right conjunct. The percentages of acceptable responses are between brackets.

(3) a. None of the astronomers saw the comet, but John did. (83%)
b. Seeing the comet was nearly impossible, but John did. (66%)
c. The comet was nearly impossible to see, but John did. (44%)
d. The comet was nearly unseeable, but John did. (17%)

[After Arregui et al. (2006)]

Although an independent grammar should abort any interpretation of ill-formed ellipsis constructions, it seems that they can be saved online. Constituting the ultimate tension between competence and performance, the question is why and when during comprehension the "parser" would overrule licensing instructions imposed by the grammar so easily, which in itself seems a licensing issue. Apparently, to develop linking hypotheses in a two-system approach one would need to account for mutually constraining factors. Within a single cognitive system that maps linear strings – sometimes incomplete (2) or sometimes (relatively) ill-formed (3) – to conceptual representations and vice versa, the attested gradience could be plausibly captured. Such a system would build representations of a grammar that amount to instructions ranging from higher-level (grammar) to lower-level (processing) procedures, theory being a kind of abstraction or idealisation of the parser (Sprouse & Almeida, 2013).

With respect to representations of antecedents of well-formed ellipsis constructions, experimental research has put forward at least two mechanisms that seem to fall on different sides of the familiar syntactic-semantic divide. Either a copy of *bought a book* in (2) (proposed as "Copy α" by Frazier & Clifton, 2001) or a more fully interpreted discourse representation that is directly accessible is inserted in the ellipsis site (implemented as a cue-based pointer mechanism by Martin & McElree, 2008). Both accounts assume that sentence comprehension is an incremental process during which incoming structural information is paired with an interpretation – updating representations step by step. In terms of retrieving and integrating the antecedent, both accounts predict, rather unhelpfully, the same behavioural results, stating that the speed of interpreting the ellipsis does not depend on antecedent complexity. For example, no difference would be observed between (2), repeated here in (4a), and (4b).

(4) a. Eva bought a book in the shop, and Agnes <*e*> in the supermarket.
  b. Eva bought a book about gardening in the shop, and Agnes <*e*> in the supermarket.

However, it is important to distinguish retrieval from integration as it seems reasonable that a copy account may be rather beneficial to an integration process, while the relative cost of searching and finding structure might increase as a function of its size. Contrastingly, a cue-based account, which is mainly explaining the mechanism of retrieval, may predict the reverse. Figure 4.1 may help to explain how both mechanisms predict an *equal* processing cost with respect to the resolution process as a *whole*.

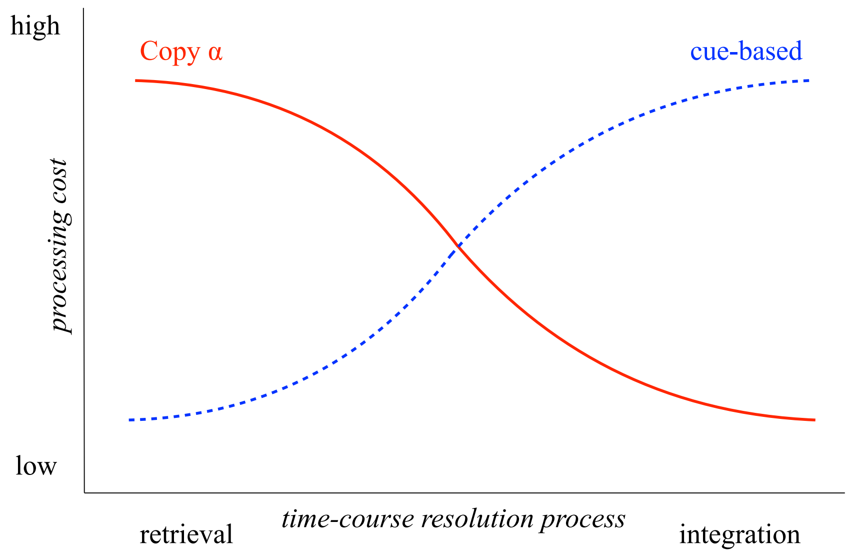


Figure 4.1: Schematic representation of predictions made by Copy α and a cue-based pointer mechanism, in terms of processing cost related to retrieval and integration processes.

To decide between the two mechanisms, the electrophysiological technique of event-related potentials (ERPs) may provide key insight, being the method of choice to investigate the time-course of cognitive processes. Effects on mechanisms of retrieval are expected early on, followed by those that impact on the integration process. In accordance with a copy mechanism, the onset of ERP signatures relating to accessing and copying missing structure would vary as a function of structure size; upon retrieval a fully-fledged structure would facilitate the integration process predicting relatively small effects. As mentioned, a "cue-based" approach would account for the reverse situation. Early ERP signatures of retrieval may be fixed since the antecedent is directly accessible. On the other hand, integration processes may operate on representations of various types, as discourse information has to be integrated in incrementally built-up structure, predicting ERP variability relatively late in the time course. Thus, ERPs may be used to compare models that are able to make predictions regarding timing. With this method, we may gain valuable insight with respect to the division of labour of syntactic, semantic and prosodic constraints. In the current study, these dimensions come under investigation – taking up the challenge of integrating theoretical conceptions of ellipsis resolution with cognitive performance data.

4.1.4 Towards a unified research program

Three levels of analysis

In the preceding sections, we have seen that a model based on a sharp distinction between static knowledge and processing mechanisms is not suitable to accommodate both theoretical and experimental research. One attempt to provide a framework that relaxes the competence-performance opposition has been put forward by Jackendoff (2002). Because this model emphasises the independent combinatorial character of syntactic, semantic and phonological information types, it seems to be particularly suitable for investigating the multidimensional character of phenomena such as ellipsis. However, his proposal is not sufficiently specific with respect to neurophysiological data such as ERPs to be truly integrative. The lack of a proper integrative theory has led (Poeppel & Embick, 2005:103) to provocatively forecast "(long-term) interdisciplinary cross-sterilisation rather than cross-fertilisation between linguistics and neurobiology, or, for that matter, linguistics and other empirical disciplines." In other words, we need a methodological framework that also incorporates physiological data. Furthermore, such a framework should specify hypotheses concerning the linking of theory and data.

Recently, Marr (1982)'s model for investigating vision has been put forward as reference to bring together linguistics and neuroscience (see for example Baggio, Lambalgen, & Hagoort, 2012; Embick & Poeppel, 2015). This model is built on three levels of analysis: (a) Computational Theory, (b) Representation and Algorithm and (c) Hardware Implementation. According to Marr any machine carrying out an information-processing task (i.e. a cognitive process) must be understood by answering the questions (a) what is computed?, (b) how is computation carried out? and (c) how can the computation be realised physically? An ideal integrated theory of language would then be a combination of formal theories of grammar, language processing and neural computation, respectively. Entertaining these levels as descriptions of one cognitive system, grammar could be understood as the abstract description of the representations that this system builds (Lewis & Phillips, 2015); representations that are identified and put together during comprehension and production, respectively. At the same time, such an integrated theory would help us to investigate explanatory connections between all three levels of analysis; for example, we could ask to what extent discoveries about the structure and functional organisation of the brain explain (rather than just describe) properties of the computations and representations that constitute language (Embick & Poeppel, 2015).

On the surface, the proposed framework may still resemble a competence-performance distinction, though with an added neurobiological level. It should be noted, though, that this system requires a theory of computation that can actually be carried out in real time. In that sense, "mentalistic" linguistic theory ought to proceed by according the same value to experimental results as it does to evidence from native speaker intuitions (Poeppel & Embick, 2005). Whereas in the "old distinction" a variety of phenomena were lumped together under the umbrella of performance, (Baggio et al., 2012:339) note that in the proposed approach these phenomena may be disentangled at a representational level and "understood in their distinctive features". For example, algorithms that specify working memory constraints may shed light on the type of data structures that a computational theory should produce, as well as the memory architecture and its neuronal substrates. Concentrating mainly on semantics, Baggio et al. stress – in line with Marr – the importance of the computational nature of integrating constraints derived from all levels of analysis.

Computational (psycho)linguistics

Hypothesis testing leads to theoretical progress. We have arrived at a model in which multiple sources of data can be taken into account and which enables us to test hypotheses locally, within levels. Preferably, they survive across levels. The less local the test domain, the more variables will have to be taken into account. The proposed integrated endeavour promotes a computational approach requiring rigid specificity which is simultaneously well-suited for testing hypotheses in highly controlled experiments. In terms of a division of computational approaches suggested by Cremers and Hijzelendoorn (2014), "gnostic", "paragnostic" and "agnostic" methods virtually align with the levels of description that we are now familiar with: formal theory, production and comprehension mechanisms and neural behaviour, ranging from "knowledgeable" (gnostic) explanatory linguistics to "naive" (agnostic) connectionist approaches in neurolinguistics. Intuitively, an integrated approach should aim at a computational model that combines symbolic and sub-symbolic terms, bridging the continuum of data collection methods and ultimately being explanatory at all levels of analysis.

In the meantime (computational) research proceeds step by step. This way, we will be able to determine the relative gnostic weight. (Lewis & Phillips, 2015:30) point to promising computational accounts that are based on transparent grammar-to-parser mappings, arguing that such models "may be understood as relating different levels of analysis, as in a one-system approach, rather than relating independent cognitive systems." The same transparency can be found in Cremers and Hijzelendoorn's ongoing project, "Delilah" (see for example Hijzelendoorn & Cremers, 2009; Reckman, 2009). Delilah is an example of a pure gnostic machine which parses and generates Dutch sentences on the basis of precise syntactic and semantic symbolic representations. Cremers and Hijzelendoorn (2014) also note the inevitability of incompleteness of grammar. The question is to what extent it can be supplemented by other terms than just symbolic ones. A "semantic machine" such as Delilah would require a semantic database – provided by computationally-based corpora research – other than the "lexikon", which is not (yet) available.

An example of a hybrid model of sentence processing is based on a well-established cognitive architecture *Adaptive Character of Thought-Rational* (ACT-R), proposed and primarily developed by John Robert Anderson (see act-r.psy.cmu.edu for a list of relevant applications and publications). Lewis and Vasishth (2005) developed a model that is able to simulate human reading time data. Utilising principles of memory retrieval and controlled processing, it is far from complete as its functionality only revolves around cue-based retrieval during syntactic parsing in the course of reading. However, it is flexible to the extent that it allows the researcher to add assumptions and theories about a specific task to be modelled. Furthermore, although a precise theory of cues is still lacking, an ACT-R based model may provide us with a tool to determine the nature of effects of interference, locality, antilocality and storage effects in sentence processing. It may turn out, to the dismay of some formal linguists, that some linguistic phenomena are grounded in principles of general cognitive processes. Yet, a means to estimate the limits of formal conditions is exactly what we need – even if one would still subscribe to the "old distinction". If we embrace a computationally sound approach, we can speak of one system that provides the representations that both listeners and speakers arrive at during language use.

4.1.5 Conclusion

I have argued that we should understand the human language system in terms of three levels of description that ultimately amount to a computational model. While I endorse the hypothesis that contradictory outcomes from offline versus online data may be due to different stages of the computations they tap into, future research is needed to confirm this. Computational linguistics may add valuable insights just as theoretical, descriptive and (other) experimental research does; furthermore, it has the benefit that it may provide us with data produced by highly controlled experiments. Testing computational models and integrating and manipulating the amount and type of predefined constraints will enable us to bridge theory and data – provided that real-time computation is a shared level of explanation. Within an integrated approach, we may determine which linguistic constraints are essentially linguistic and which of them are manifestations of more general cognitive capacities. A platform is at hand to overcome the persistent gnostic-agnostic divide.

4.2 Working memory load

Over the last few decades, the period in which neurobiological research has expanded dramatically, it has become clear that no human brain is the same. Although the gross (functional) anatomy seems to be similar among the population to some degree, there are certainly individual differences. If we assume that every human employs automatic linguistic processes, we may not expect to see differences in this area. However, in the field of language comprehension individual differences have indeed been found (see for example Kaan et al., 2013; Otten & Van Berkum, 2009). Such variation may lead to differences in (amplitudes of) ERP components and may be ascribed to variation of the capacity of people's working memory systems. This is expected to play a major role during ellipsis processing. To control for such variability, I will test participants by means of an additional memory task. This section explains the nature of this task and why we have chosen to use it.

4.2.1 A model of working memory

The term 'working memory' (WM) stems from the earlier proposed notion of short-term memory (STM). Following Baddeley (see for a history and overview Baddeley, 2012), I will regard STM as a system for "simple temporary storage of information, in contrast to WM, which implies a combination of storage and manipulation." It is understood by Baddeley as a multicomponential capacity comprising four subsystems:

- a phonological loop, concerned with verbal and acoustic information
- a visuospatial sketchpad, the visual equivalent of the phonological loop
- an episodic buffer, a multi-dimensional buffer store that links between WM components, but also links WM to perception and long term memory
- a central executive system, an attentionally-limited system which links to the episodic buffer

WM, as a whole, serves the function of integrating the information types that are processed by the phonological loop and visuospatial sketchpad into a unified representation. This representation may be stored for a short while and be manipulated upon. With regard to language comprehension, the episodic buffer is of particular interest: it has been suggested that its capacity may predict the aptitude of prose comprehension, as we will see below.

4.2.2 Working memory and sentence comprehension

The seminal study by Daneman and Carpenter (1980) suggests that there is a correlation between, in their terms, "WM span" and the capability for prose comprehension. Subsequent research, using paradigms where participants were required to employ a combination of temporary storage and processing, corroborated Daneman and Carpenter's findings (see for a meta-analysis Daneman & Merikle, 1996). Daneman and Carpenter (1980)'s paradigm – in which participants read out a series of sentences of different lengths while having to remember the last word in each sentence – has since become classic.

It has further been suggested that a separate subsystem of the WM is employed to assign syntactic structure to a sentence and to use this structure to determine the meaning of that sentence (e.g. Caplan & Waters, 1999). Caplan and Waters suggest that WM capacity as a whole might not be associated with differences in the efficiency of syntactic processing in sentence comprehension. Rather, the process of "recognizing words and appreciating their meanings and syntactic features; constructing syntactic and prosodic representations; and assigning thematic roles, focus, and other aspects of propositional and discourse-level semantics" (Caplan & Waters, 1999:78) might call on a different pool of resources. They propose a separate sentence interpretation resource theory, which assumes that general WM tasks cannot be used to predict language processing efficiency. We could interpret this standpoint as referring to the hypothesis that automatic linguistic processes are independent of general WM. However, this study was written before the construct of the *episodic buffer* had been put forward. This module was proposed to account for the fact that an executive system should be able to link to a temporary storage. Daneman and Carpenter (1980) and follow-up research had shown that such storage should be bigger than the limited capacities of the phonological loop and visuospatial sketchpad. The function of the buffer should be to integrate and maintain information into coherent episodes. Further, a proper link could now be established between WM and long-term memory. In sum, the episodic buffer is assumed to play a major role in "binding information from diverse sources into unified chunks" (Baddeley, 2007:148). In a commentary on Caplan and Waters (1999), Kane, Conway, and Engle (1999:102) note that: "[...] working memory capacity is needed only under attention-demanding circumstances, and, insofar as syntactic processing appears to be immune to divided-attention conditions, it likely occurs relatively automatically." While syntactic aspects of language comprehension may be carried out automatically, the episodic buffer may enable us to explain how relatively more demanding tasks in which information needs to be stored (such as in ellipsis resolution) are executed.

From a neurophysiological perspective, some authors have proposed that some capacity of the human brain might be specifically devoted to syntactic working memory which appears to be a "bilateral network of inferior frontal and superior temporal brain regions, with a left lateralisation within the inferior portion of the pars opercularis of the left inferior frontal gyrus (Brodmann Area 44) " (Fiebach, Schlesewsky, Lohmann, Cramon, & Friederici, 2005; Fiebach, Schlesewsky, & Friederici, 2001). These brain sites, then, could be understood as the neural basis for the subsystem as proposed by Caplan and Waters (1999). Tasks that require maintenance of information and more computation on that information (i.e. using the episodic buffer) have consistently activated the mid-dorsolateral frontal lobe, that is Brodmann Areas 46 and 9 (see for example Petrides, Alivisatos, Meyer, & Evans, 1993). The mid-dorsolateral region is believed to keep track of our thoughts and memories; indeed, poor maintenance and manipulation of information is associated with impaired dorsolateral regions (e.g. Cannon et al., 2005).

4.2.3 Testing working memory

Verbal WM capacity can be tested by means of reading span (as introduced above). In the spirit of Daneman and Carpenter (1980), Noort, Bosch, Haverkort, and Hugdahl (2008) designed a reading span test that is compatible across four languages. In this computer-administered test, five trials of 20 sentences are presented. However, note that this task might be too demanding in an experiment where participants have already carried out a sentence reading or listening task beforehand; therefore I will avoid this situation in my own experiments. Besides, span tests are regarded by Caplan and Waters (1999) as calling on another resource than syntactic processing. It has been shown that the mid-dorsolateral region is implicated in the monitoring and manipulation of information in working memory (Petrides, 2000). Ideally, I should use a design in which a stimulus needs to be temporarily stored and be recalled after intervening structure has been processed. Therefore I am looking for a method that assesses WM storage and processing in a relatively short time. Petrides et al. (1993) may offer such a method, which can be understood as a variant of Daneman and Carpenter (1980). Since Petrides et al. and his colleagues were using Positron Emission Tomography as measure (a procedure during which participants are injected with a short-lived radioactive substance), they wanted to minimise the scanning period. Hence, the design was compact. Below, I reproduce their description of the testing procedure in which participants carried out "self-ordered" and "externally ordered" number generation tasks .

> [...] the subjects were scanned with PET for 60 sec under three different conditions of testing. In the control condition, the subjects were required to count aloud from 1 to 10 at the rate of approximately one digit per second. They were told that when they reached the number 10, they were once again to start counting from 1 to 10 and continue in this manner until told to stop. In the self-ordered condition, the subjects were asked to say aloud, in a random order, the numbers from 1 to 10. They were asked to monitor carefully the numbers they gave so as not to repeat the same number more than once until all 10 numbers were reported. At that point they were to begin a new trial (i.e., a sequence), again generating numbers randomly from 1 to 10. The subjects were asked to start always from the number 1, because this would permit the experimenter, who was recording the responses, to know when a new trial had begun. As in the control condition, the subjects were told to generate the numbers at the rate of approximately one per second. An average of 5.25 trials (range, 4.5-6.0) was completed during scanning, with an average error of 0.9. An error was defined as a repetition or an omission of a number in a trial. In the externally ordered condition, the subjects were told that, during scanning, the experimenter would read out in a random sequence the numbers from 1

> to 10, omitting one of these numbers. The subjects had to monitor carefully the numbers read by the experimenter because, on completion, they would have to say the number that had been omitted. The experimenter would then administer another trial - i.e., read another random sequence of the numbers 1 to 10, again omitting one number that the subject would be required to report. The numbers were read out at the rate of approximately one digit per second. An average of 5.6 (range, 5.0-6.0) trials was completed during scanning and the subjects made an average error of 0.2 per trial.
>
> Before each scanning condition, the experimenter explained the requirements of the task to be performed and the subjects practiced the task once. The subjects kept their eyes open during scanning, but visual stimulation was reduced by dimming the lights within the scanning room and by surrounding the subject with black curtains.
>
> [Petrides et al. (1993:880)]

The advantages of this paradigm are that the performance of the participants can be related to a fixed control condition and the critical conditions are related to the mid-dorsolateral region. Furthermore, it is a task in which participants have to fill a gap in some sequence, a procedure that is at least in part reminiscent of the resolution of ellipsis in language. In addition, it takes at most ten minutes to complete.

4.3 Hypotheses and possible results

We have arrived at a comparison of two behaviourally motivated models of ellipsis processing that reflect the syntax-semantics divide in the theoretical literature to a certain extent, namely Copy α and a cue-based pointer mechanism. Despite this, it has also become clear that a mapping between existing theoretical insights and neurophysiological processes may not always be straightforward or even justifiable. Therefore, although it is my intention to integrate theoretical approaches with the investigation of processes at a neural level, the results of the current study should nonetheless be interpreted with great caution when trying to relate the findings to theoretical notions.

Each of the experiments in this study aims to investigate the online time course of the processing of Gapping and Stripping and to what extent it may be modulated by syntactic, semantic and prosodic factors. In Chapters 6, 7, and 8, I report ERP studies testing syntactic, semantic and prosodic variables respectively. Let us consider overall hypotheses based on the theoretical discussions presented in the foregoing chapters.

Effects of the manipulated variables on mechanisms of retrieval are expected early on, followed by those that impact on the integration process. A copy account predicts modulation of ERP signatures related to syntactic processes early in the time course, possibly manifesting as (E)LAN effects. The relative cost of searching and finding structure might increase as a function of the structure's size. Because this account proposes that a full structure is available, integration processes would be carried out with relative ease. Contrastingly, a cue-based account, which is mainly explaining the mechanism of retrieval, would predict the reverse. This account, would predict a burden on integration processes. Early ERP signatures of retrieval may be fixed since the antecedent is directly accessible. On the other hand, integration processes may operate on representations of various types, as discourse information has to be integrated in incrementally built-up structure, predicting ERP variability relatively late in the time course. Therefore, modulation of cues would presumably be reflected by a modulation of P600 effects. A caveat is in order here, as it is not always clear what exactly can be considered to be cues. I take it that a cue can be related to any information type that is stored in a more fully interpreted chunk: syntactic, semantic and prosodic. Since a cue is directly accessible during processes of retrieval, it is expected that ERP signals related to retrieval are relatively small. In addition, I hypothesise that the processor may exploit a composite of different cues, if needed. I have no prediction, however, as to the nature of a possible ERP signature (or signatures) that might be implicated therein. As discussed, there is not much ERP research on Gapping, so this study must be regarded as somewhat exploratory. While this limits the way I can 'stand on the shoulders of giants', I think it is important nonetheless to do such studies.

Before proceeding to report the ERP experiments in Chapters 6-8, it is first necessary to describe the preparatory tests for these experiments; this is the topic of the next chapter.

CHAPTER 5

Replication and norming stimuli

Essential to the process of science is the replication of previous studies in order to validate existing findings before building on them. As has become clear in Chapter 3.5, electrophysiological research on ellipsis, let alone Gapping in Dutch, is scarce. Section 5.1 reports the findings of a replication of Kaan et al. (2013). I thank Wouter Broos for his assistance with organising the stimuli and recording of the EEG data. In sections 5.2 and 5.3 I report norming studies that I carried out to pretest newly designed stimuli to be used in subsequent experiments.

5.1 Validating ERP results: a replication study

5.1.1 Method

Test materials

Using a Latin Square design, 117 quadruplets as described in chapter 3.5 were divided over four lists and complemented with 96 fillers.[1] The experimental paradigm is illustrated again in (1). The four stimulus conditions and additionally the two collapsed conditions (No-Gapping vs. Gapping) are colour-coded, corresponding to colours used in graphs later on.

(1) a. Anouk zond de kaart aan haar vader, en Julia **de** **bloemen** aan haar moeder.
Anouk sent the card to her father, and Julia the flowers to her mother.
'Anouk sent the card to her father,
and Julia the flowers to her mother.' (***Plausible Gapping***)

b. Anouk schreef de kaart aan haar vader, en Julia **de** **bloemen** aan haar moeder
Anouk wrote the card to her father, and Julia the flowers to her mother.
'Anouk wrote the card to her father,
and Julia the flowers to her mother.' (***Implausible Gapping***)

c. Anouk zond de kaart aan haar vader, terwijl Julia **de** **bloemen** aan haar moeder stuurde.
Anouk sent the card to her father, while Julia the flowers to her mother shipped.
'Anouk sent the card to her father,
while Julia shipped the flowers to her mother.' (***Control for condition a***)

d. Anouk schreef de kaart aan haar vader, terwijl Julia **de** **bloemen** aan haar moeder stuurde.
Anouk wrote the card to her father, while Julia the flowers to her mother shipped.
'Anouk wrote the card to her father,
while Julia shipped the flowers to her mother.' (***Control for condition b***)

[Kaan et al. (2013)]

[1] The odd number has to do with the fact that, in the original experiment, from the original set of 120 items three had been omitted due to an experimenter error.

To recapitulate, in Kaan et al. (2013), at the determiner, a LAN effect was expected for Gapping conditions ***a-b*** versus No-Gapping conditions ***c-d***, but this was only apparent in a group of participants who scored relatively poorly at the end of sentence task. At the noun following the determiner, an N400 for (***b*** versus ***a***), perhaps followed by a P600, was predicted. Only a P600 effect turned out to be significant; an N400 was arguably detected, but only as a numerical trend. Finally, the authors hypothesised that if syntactic integration is more effortful in Gapping versus No-Gapping constructions, a P600 effect for Gapping versus No-Gapping constructions at the noun should be found. A notable result of the original study is that this effect was apparent for plausible conditions ***a*** versus ***c***.

Participants

Twenty-two native speakers of Dutch with normal or corrected-to-normal vision participated. All participants reported not to have any neurological problems or disease. Due to bad signal (3 participants) and left-handedness (1 participant), four participants were discarded and the analysis below is based on 18 right-handed participants (16 women, 2 men, $M_{\text{Age}} = 23.17$, range 20-27). Participants gave informed consent before the study and were paid €15. The experiment complied with the Ethics Committee regulations of the Faculty of Social Sciences of Leiden University, which approved its implementation.

Procedure

Participants were comfortably seated in a dimly lit sound-proof room at a distance of approximately 80 cm of a 17 inch CRT monitor. Sentences were presented one word at a time in white letters in Verdana font (18pt) on a black screen using the presentation software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Each sentence was preceded by a fixation cross ("+") which appeared at the centre of the screen and remained there for 1,000 ms. Then, each word was presented for 300 ms, followed by a 200 ms blank screen. A word before a comma was presented with that comma appended; similarly, the last word of each sentence was marked with a full stop. 1,500 ms after offset of the sentence-final word a prompt, *OK of SLECHT* ("OK or BAD"), appeared. The left response button was linked to *OK* and the right one to *SLECHT*. As a means of counterbalancing, half of the participants received the prompt and button choices the other way around. After a response click, a blank screen appeared for 1,000 ms. After every 12 sentences, participants were offered a break. Before starting the experimental phase, six warm-up practice trials were presented to the participants. These sentences bore no resemblance to any of the experimental or filler items.

In addition to the experiment reported above, a working memory test based on a task described in chapter 4.2.3 was carried out. Participants were instructed to count aloud from 1 to 10 at the rate of approximately one digit per second (5 trials). In the second session, they were asked to randomly count aloud numbers from 1 to 10 while monitoring that every number was only mentioned once in each trial: they were not allowed to repeat the same number more than once until all 10 numbers were reported. During the third session participants listened to a random sequence of nine digits between 1 and 10, after which they were asked to say which digit between 1 and 10 had been omitted (5 trials). The last session was as the third session but instead the numbers were presented visually one by one. The working memory test was carried out after the EEG recordings.

The experiment took about 2 hours per participant in total, including set-up.

Apparatus and electrophysiological recording

The electroencephalogram (EEG) was obtained using BioSemi ActiveTwo system (BioSemi B.V., Amsterdam) following the international 10/20 system (originally proposed by Jasper in 1958 and, after modifications, standardised as of 1991 by the American Electroencephalographic Society) Ag/AgCl electrodes (Fp1/2, FC5/6, AF3/4, Fz, CP5/6, CP1/2, Cz, F7/8, F3/4, T7/8, C3/4, Pz, FC1/2, P3/4, O1/2, Oz, P7/8, PO3/4). Four flat electrodes were used to monitor the eye movements (i.e. to obtain an electro-oculogram or EOG): two above and underneath the left eye to measure blinks; two at the external canthi of both eyes to measure saccades. A flat electrode was placed on each mastoid to be used for off-line re-referencing. The EEG signal was recorded using the BioSemi ActiView software at a sampling rate of 512 Hz. Electrode impedance was monitored during installation and running to ensure a low level of noise.

Data analysis

Using Brain Vision Analyzer Version 2.0 (Brain Products, Munich, Germany), the EEG data were preprocessed before analysis to reduce noise and artifacts as much as possible. EOG artifacts were corrected using the Gratton, Coles, and Donchin (1983) algorithm. Remaining artifacts were rejected and checked visually on the basis of the following criteria: the maximum allowed voltage step was 20 μV/ms, the maximal allowed difference of values was 100 μV in an interval of 200 ms and the lowest allowed activity was 0.5μV. Just as in the original study, a low cutoff filter of 0.16 Hz, 24 bB/oct and a high cutoff filter of 30 Hz, 24 dB/oct were applied. Epochs of 1,300 ms were computed with a 100 ms pre-stimulus baseline. ERP grand averages were time-locked to (*i*) the critical determiner following the position of the elided verb (average percentage rejected: 24.41% of the trials for Gapping and 25.74% No-Gapping conditions) and (*ii*) the noun following the determiner (average percentage rejected: 24.22% for Gapping and 24.60% for No-Gapping).

Again in accordance with the original study, the effect of Gapping versus No-Gapping at the determiner was analysed using the mean amplitude in the

100-200 ms (ELAN) and 400-600 ms (LAN) time windows. An additional time window of 200-400 ms was taken into account. At the following noun, the mean amplitude in the 300-500 ms (N400), 500-700 ms, 700-900 ms (P600), and 900-1,200 ms time windows (late positivity) were analysed.

Analyses were conducted separately for midline sites (Fz, Cz, Pz) and for the lateral electrode regions: left/right frontal (Fp1/2, AF3/4, F7/8, F3/4), left/right central (FC5/6, T7/8, C3/4, CP5/6), left/right parietal (P7/8, P3/4,PO3/4, O1/2). For each time window, a repeated measures analysis was carried out with as within-subjects factors GAPPING, ANTERIORITY (3 levels), and, for analyses involving lateral sites, HEMISPHERE (2 levels). Additionally, for the epochs of the noun position, PLAUSIBILITY of the verb in the first clause and object in the second. Mean voltage-amplitude was considered as the dependent variable in the analysis, and p-values where corrected for sphericity where required.

Throughout this thesis, both the behavioural data and the electrophysiological data were analysed using **R** version 3.3.3 (R Development Core Team, 2008). As can be seen above, I use small capitals to indicate factors (variables). Scripts and data can be found at http://bobbyruijgrok.com/data.

5.1.2 Behavioural results

Average accuracy rates of the acceptabilty judgements were high and no participants were rejected on the basis of accuracy ($M = 87.45\%$, $SE = 0.96\%$). The accuracy scores were similar across conditions ($M_{\text{Plausible Gapping}} = 88.82\%$, $M_{\text{Implausible Gapping}} = 87.70\%$, $M_{\text{Plausible control for a}} = 86.70\%$, $M_{\text{Control for b}} = 86.21\%$). The difference in mean values was not significant as shown by a repeated-measures ANOVA by participants with CONDITION as independent factor and ACCURACY OF SENTENCE COMPREHENSION as dependent variable [$F(3, 51) = 0.32, p = .808$, $\eta^2_G = .010$].[2]

Although the working memory task consisted of four sessions, I will only report the findings of the last three: the first session was meant as a control condition as to adjust the participant's speed of production to one digit per second approximately. Errors were defined as follows: a repetition or an omission of a number in a trial of the self-oredered condition, or an incorrect response in a trial in the auditory and visual conditions. The accuracy ratio of the three test sessions was 67.04% ($SE = 2.87\%$). Per condition the scores were: $M_{\text{Random Counting}} = 66.67\%$, $M_{\text{Auditory Presentation}} = 58.89\%$, $M_{\text{Visual Presentation}} = 75.56\%$. Although numerically the difference between the auditory and visual conditions seemed large, a repeated measures ANOVA by subjects with CONDITION as independent factor and ACCURACY OF NUMBER RECALL as dependent variable yielded only marginal significance [F(2,

[2]Throughout this thesis, in reporting repeated measures ANOVAs I report the generalized eta squared as is proposed by Bakeman (2005) as useful statistic: .02 = small, .13 = medium and .26 = large.

34) = 2.72, p = .080, η^2_G = .084].

The scores of the comprehension task of the ERP experiment were compared with the scores of the working memory task. A slight correlation was found between the variables but this was not statistically significant ACCURACY OF SENTENCE COMPREHENSION and ACCURACY OF NUMBER RECALL [r = .389, p = .110].

5.1.3 Electrophysiological results

ERPs at the determiner

Figure 5.1 shows the ERPs for the Gapping and No-Gapping conditions (i.e. collapsed over plausibility conditions: ***a-b*** and ***c-d***) at the moment the critical determiner was displayed. Relative to No-Gapping conditions a negativity can be observed in the Gapping conditions starting just after 200 ms at all electrodes.

On midline electrodes, the factor GAPPING reached marginal significance in the time window 200-400 ms post-onset [F(1, 17) = 3.44, p = .081, η^2_G = .022]. No other effects could be established.

On lateral electrodes, the factor GAPPING reached significance in the 200-400 ms time window [F(1, 17) = 5.33, p = .034, η^2_G = .018] as well as the 400-600 ms time window [F(1, 17) = 6.01, p = .025, η^2_G = .023]. In the 100-200 ms time window the factor HEMISPHERE yielded significant effects, the left-lateralised electrodes having more negative averaged amplitudes [F(1, 17) = 11.22, p = .004, η^2_G = .042]. Significant effects of HEMISPHERE coincided with significant interaction effects of ANTERIORITY by HEMISPHERE in the 200-400 ms [F(2, 34) = 5.07, p = .012, η^2_G = .007] and 400-600 ms time window [F(2, 34) = 4.23, p = .023, η^2_G = .005]. The interaction effects are visualised in Figure 5.2. As can be seen, left central electrodes show relatively negative mean amplitudes.

To investigate whether the overall effect was attenuated by individual variation, the mean differences in amplitude in all three time windows between the Gapping and No-Gapping conditions, collapsed over the left anterior electrodes (Fp1, AF3, F7, F3), were analysed with respect to (*i*) sentence judgement accuracy of the experimental items and (*ii*) the accuracy of the working memory task. No significant correlations could be established.

ERPs at the noun

Effects of semantic integration between the noun and the elided verb were first analysed in relation to the factor PLAUSIBILITY. ERPs at the critical noun are displayed in Figure 5.3 for the Plausible Gapping and Implausible Gapping conditions (***a*** and ***c***).

While a negative deflection can be observed at around 400 ms, no significant effect of PLAUSIBILITY could be established in the 300-500 ms time win-

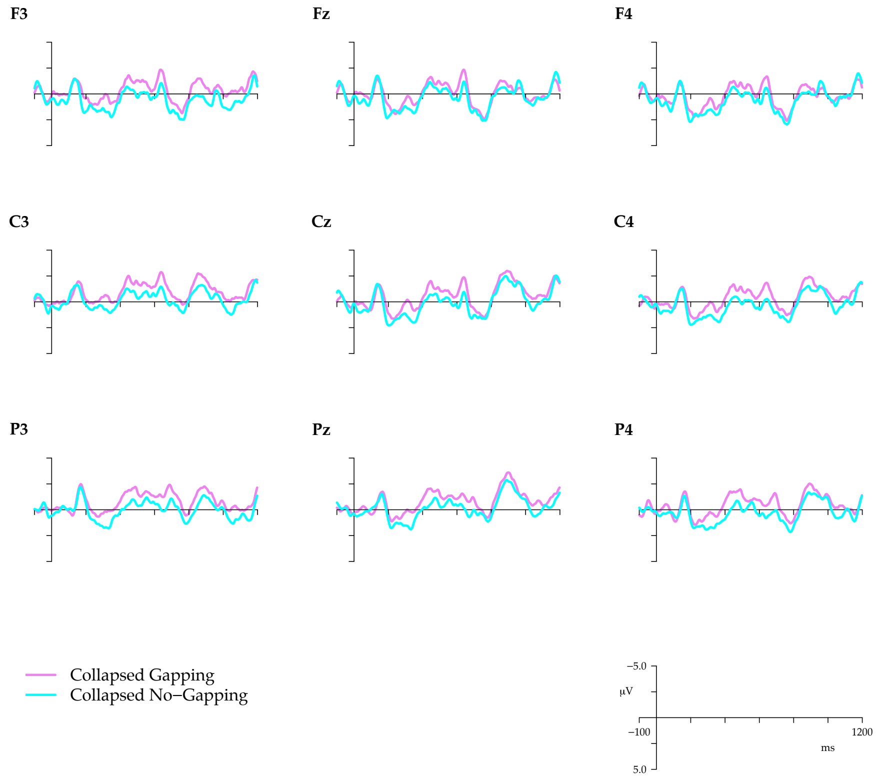


Figure 5.1: Grand averages of collapsed Gapping conditions (***a*** and ***b***) compared to No-Gapping conditions (***c*** and ***d***) at onset (y-axis) of the determiner (*de*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 86.

dow. However, on midline electrodes, an effect of ANTERIORITY was apparent [$F(2, 34) = 8.02$, $p = .007$, $\eta^2_G = .063$]. In the same time window on lateral sites an effect of HEMISPHERE could be observed [$F(1, 17) = 10.57$, $p = .005$, $\eta^2_G = .029$]. These effects were due to relatively negative amplitudes at right-lateralised centro-parietal sites.

In Figure 5.3 a late positivity for Implausible Gapping can be observed most prominently at electrode Pz. While no significant effects for the factor PLAUSIBILITY were found in later time windows (after 500 ms), on midline sites an effect of ANTERIORITY was established in the 500-700 ms time window [$F(2, 34) = 5.12$, $p = .022$, $\eta^2_G = .023$] and 700-900 ms window [$F(2, 34) = 8.68$, $p = .007$, $\eta^2_G = .004$]. Again, these effects were due to relative negative amplitudes at centro-parietal sites. In the 700-900 ms window on lateral

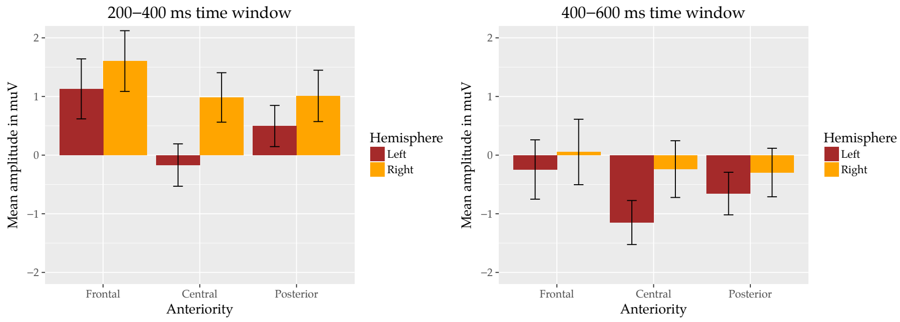


Figure 5.2: Error bar graphs of interaction effects of ANTERIORITY by HEMISPHERE at the determiner (*de*) on lateral electrodes in 200-400 ms and 400-600 ms time windows.

sites, an additional interaction effect of PLAUSIBILITY by HEMISPHERE was found [$F(1, 17) = 4.68, p = .045, \eta^2_G = .002$]. Figure 5.4 shows that implausible items caused counter effects on the mean amplitude in relation to left and right electrodes, the left hemisphere being implicated in relatively large negativity.

To further analyse integration effects of the elided verb at the position of the noun, the factor GAPPING was taken into account. In Figure 5.5, the difference between Plausible Gapping and Plausible No-Gapping conditions (***a*** and ***c***) are displayed. Relative to No-Gapping a large positive deflection can be observed for the Gapping condition.

On midline electrodes, an effect of GAPPING was found in the 700-900 ms window [$F(1, 17) = 6.56, p = .020, \eta^2_G = .037$] and in the 900-1,200 ms window [$F(1, 17) = 6.40, p = .022, \eta^2_G = .034$].

No effect of GAPPING could be established on lateral sites.

5.1.4 Discussion

In contrast to the original study, a negativity could be demonstrated at the determiner as the ERPs show an (E)LAN-like effect. This was hypothesised as a possible outcome. The interaction of ANTERIORITY by HEMISPHERE in later time windows can be explained by the relative negative amplitudes at central sites orientated at the left. In that sense, the negative component has a relatively central distribution in this study. Considering that the factor GAPPING was most prominent in the 200-400 ms and 400-600 ms time windows, the component looks like a LAN rather than an ELAN. Crucially, the effect of GAPPING was not attenuated by individual variation, yet might indeed be considered as indexing prediction processes (as was suggested in the original study). Although Gapping and No-Gapping conditions were balanced across experimental items, Gapping sentences in this study were in fact in the minority if one takes all stimuli, including fillers, into account. Of the 96 filler items,

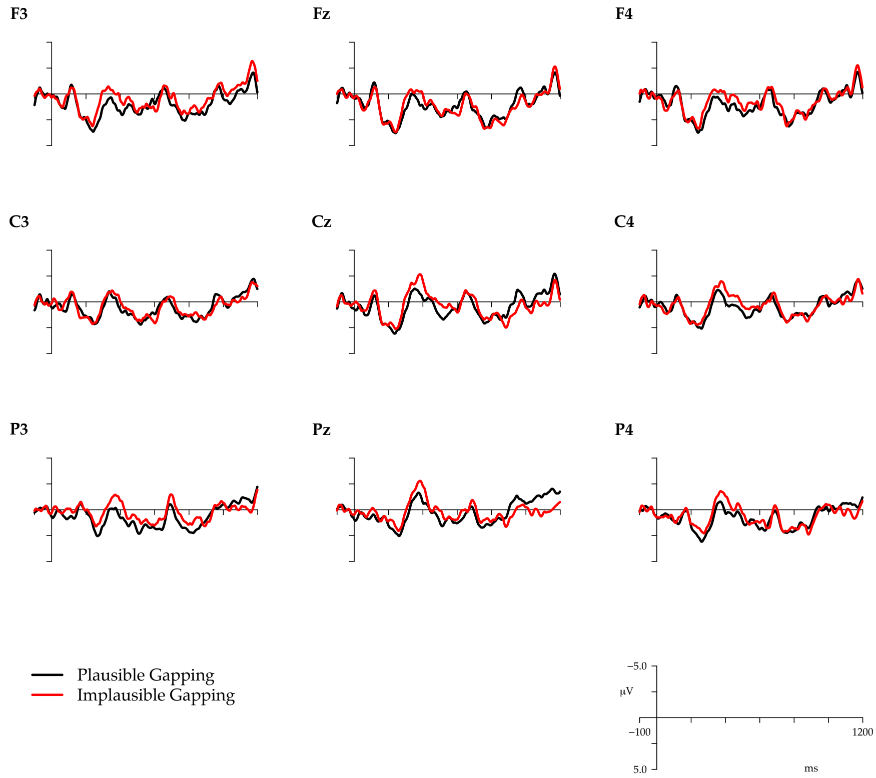


Figure 5.3: Grand averages of Plausible Gapping condition (***a***) and Implausible Gapping condition (***b***) at onset (y-axis) of the noun (*bloemen*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 86.

only 16 contained Gapping constructions, notably containing a coordination with the connective "maar".

In line with the original study, the factor PLAUSIBILITY did not yield an N400 effect at the position of the noun. Although it was numerically apparent, it was not statistically significant. Possibly, a time window of 200 ms is too large, meaning that an N400 component in this design might be expressed at a shorter latency.

The P600 effect for the factor PLAUSIBILITY in the original study could not be corroborated in this replication. A late positive deflection was visible but it was not statistically significant. The interaction effect of PLAUSIBILITY with HEMISPHERE shows that implausible items yielded opposing mean amplitudes – negative in the left hemisphere and positive in the right hemisphere.

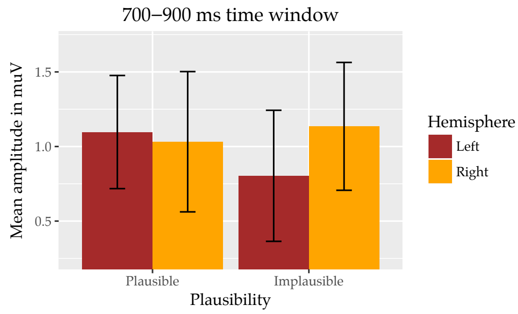


Figure 5.4: Error bar graph of interaction effect of PLAUSIBILITY by HEMISPHERE on the noun (*bloemen*) on lateral electrodes in the 700-900 time window.

This interaction may explain why no straightforward P600 could be established.

Of the most interest in relation to this thesis is the effect of the factor GAPPING, which could be corroborated in the 700-900 ms and 900-1,200 ms time windows. At the position of the noun, a process of integration may be assumed and it seems likely that this is expressed by late positive P600-like deflections. In addition, a close look at Figure 5.5 points the attention to earlier time points. It seems that a positivity is already apparent at an early stage at around 350 ms. Again, it could be that analyses using shorter time windows may have revealed significant effects here.

A few caveats are in order though. Firstly, negative deflections observed at the determiner may have had the effect of amplifying any positive effect in the epochs of the noun. Pre-stimulus activity may be problematic for the evaluation of critical time points (Luck, 2014:256). In that sense, a positivity could be seen as artefactual effect. Future designs should overcome this problem. Secondly, the analysis of this replication is based on 18 participants instead of 30 in the original study, which yields less statistical power. Nevertheless, the effect sizes for the effect of GAPPING on midline sites in the 700-900 ms and 900-1,200 ms time windows are relatively large.

5.1.5 Conclusion

In addition to an evaluation of previous studies a proper study should commence with an attempt to replicate previous published findings. Unfortunately, this prerequisite is generally seen as an unrewarding task and therefore often left out. Although results of a replication study may deviate from the original, they may still give insight as to how to proceed. The current replication gave rise to a result that was hypothesised, but which was not apparent in the original study. A LAN-like component was found that can be regarded

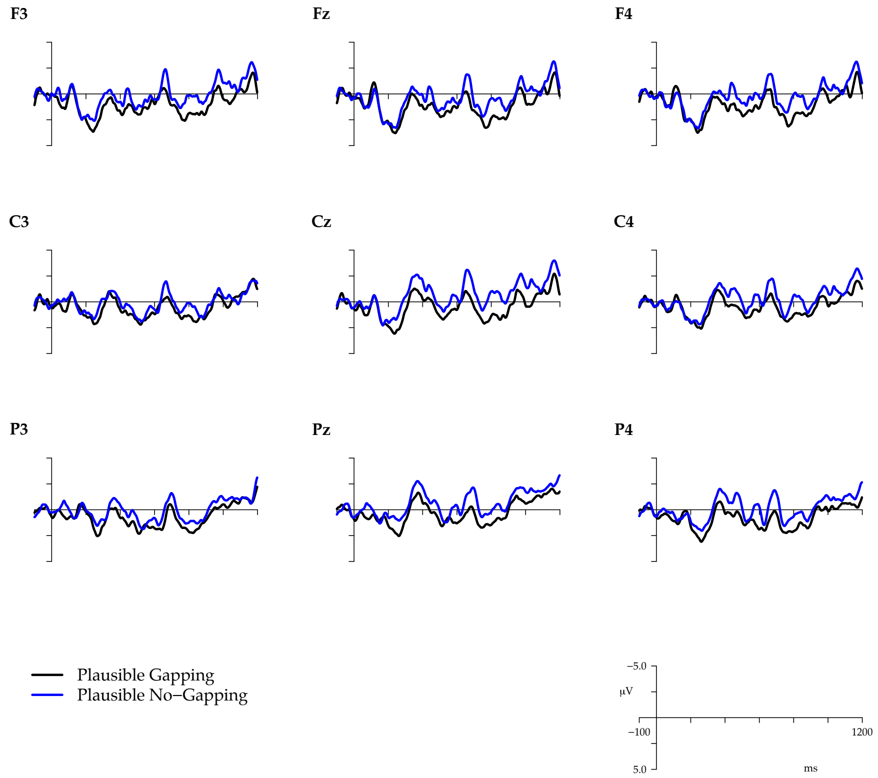


Figure 5.5: Grand averages of Plausible Gapping condition (***a***) and Plausible No-Gapping condition (***c***) at onset (y-axis) of the noun (*bloemen*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 86.

as an index of prediction. This effect seems marginally sensitive to individual variation, but may in fact be due to the relative frequency of Gapping items in the stimulus list. Furthermore, the effects of implausible items appeared to be less strong than in the original study. Again, no N400 was found and additionally a P600 was only numerically visible. However, the replication does corroborate processes of integration of a plausible elided verb at the critical noun, interpreted in the original study as being on a par with the integration of object *wh*-phrases. Gapping of plausible phrases, then, seems to be most appropriate to investigate further and this will be pursued in the continuation of the current research.

5.2 Norming stimuli I: acceptability of structural elision

5.2.1 Purpose

Throughout the experiments reported in this thesis I make use of sentences as stimuli. Preferably the stimuli should be designed such that they can be used in different experiments. This method allows us to compare results from different experiments. Furthermore, I wish to use grammatical and interpretable stimuli to investigate Gapping and Stripping. During the ERP experiments, participants will answer a comprehension question after every stimulus. On the one hand, I can make sure that participants actually *read* the sentences, on the other, comprehension scores can be analysed with respect to the complexity of the ellipsis.

Test sentences should be minimal pairs. Given that during the ERP experiments sentences will be presented by means of a word by word reading task, a fixed measure point – one word – is required to compare effects of ellipsis between conditions.This section is a report of a pilot study of stimuli in which structural complexity of the ellipsis was manipulated: phrases are cut off step by step (condition by condition) reducing the amount of overt structure step by step. The goal of this pretest is to ascertain the acceptability of the stimulus sentences, in order to be able to reject uninterpretable stimuli and gain awareness of acceptability differences across the stimuli set. The ERP experiments described in Chapter 7 were designed on the basis of the tested items.

5.2.2 Method

Building on previously used materials

Since only one peer-reviewed ERP study of Gapping processing in Dutch (the replicated study reported in section 5.1 above) had been published at the time I started this research project, it seemed most practical to develop stimuli on the basis of the test sentences from that study. As a first step, I designed 44 quadruplets as exemplified in (2).

(2) a. Omdat Hilde in de voortuin het gazon onderhield en Ralph in
Because Hilde in the front.garden the lawn maintained and Ralph in
de achtertuin de paden harkte, waren de buurtgenoten vrolijk.
the back.garden the paths raked, were the neigbours happy
'Because Hilde maintained the lawn in the front garden and Ralph raked the paths in the back garden, the neighbours were happy.'

b. Omdat Hilde in de voortuin het gazon onderhield en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin de paden, waren de buurtgenoten vrolijk.
in the back.garden the paths, were the neigbours happy
'Because Hilde maintained the lawn in the front garden and Ralph the paths in the back garden, the neighbours were happy.'

c. Omdat Hilde in de voortuin het gazon onderhield en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin, waren de buurtgenoten vrolijk.
in the back.garden, were the neigbours happy
'Because Hilde maintained the lawn in the front garden and Ralph in the back garden, the neighbours were happy.'

d. Omdat Hilde in de voortuin het gazon onderhield en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
ook, waren de buurtgenoten vrolijk.
too, were the neigbours happy
'Because Hilde maintained the lawn in the front garden and Ralph too, the neighbours were happy.'

Condition *a* represents the control sentence: a fully-fledged structure with all phrases in place. In condition *b*, the verb is elided in the right conjunct, in condition *c* the verb with the object are elided, and in condition *d* every phrase in the right conjunct except for the subject is stripped and replaced by 'too'. While the original sentences are made of conjunctions, in the new stimuli a conjunction is captured within a subordinate adjunct. The motivation to do so, was to be able to cut off phrases step by step while having a stable measuring point: the main verb *waren*. At this point, the ellipsis should be resolved. Furthermore, the completion of the subordinate clause does not hinge on the main clause as would be the case with a subject clause (e.g. *That John bought a book surprised his mother.*). In such sentences the main verb needs the subject – the whole subordinate clause – in order to integrate the arguments. As a consequence, this process may overshadow the ellipsis resolution mechanism. As can be seen in (2), the stimuli are closely related to the crucial stimuli as used by Kaan et al. (2013), repeated here in (3).

(3) a. Hilde onderhield het gazon in de voortuin en Ralph de
Hilde maintained the lawn in the front.garden and Ralph the
paden in de achtertuin.
paths in the back.garden
'Hilde maintained the lawn in the front garden and Ralph the paths in the back garden.'

b. Hilde onderhield het gazon in de voortuin terwijl Ralph de
Hilde maintained the lawn in the front.garden while Ralph the
paden in de achtertuin harkte.
paths in the back.garden raked
'Hilde maintained the lawn in the front garden while Ralph raked the paths in the back garden.'

As discussed in Chapter 3.5 Kaan et al., compared (3a) with (3b), which differ in structure. By contrast, my aim is to compare measurement point(s) between sentences with the same structure. Kaan et al. utilised the noun phrase *de paden* as measuring point. Note that in (3a), this phrase is in a main clause, while in (3b) it is in a subordinate clause. As explained earlier, they reasoned that the determiner is expected in (3b) and not expected in (3a). In that sense, their results are contingent on expectancy effects which are partly induced by the clause type, i.e. the conjunction.

Only 44 stimuli could be used of the available 117 from Kaan et al since some of their original stimuli contained noun phrase modifiers. A disadvantage of such sentences for the purpose of cutting off phrases step by step is, that such modifiers cannot appear on their own and hence cannot be used in the proposed setting. For example, in (4) *de staking* cannot be separated from *van de monteurs*. This problem does not arise with adjuncts as is shown in (5).

(4) a. Renate organiseerde de staking van de monteurs.
Renate organised the strike of the mechanics
'Renate organised the strike of the mechanics.'

b. *Van de monteurs organiseerde Renate de staking.
of the mechanics organised Renate the strike
int: 'Of the mechanics Renate organised the strike.'

(5) a. Renate organiseerde de staking in de ochtend.
Renate organised the strike in the morning
'Renate organised the strike of the mechanics.'

b. In de ochtend organiseerde Renate de staking.
in the morning organised Renate the strike
'In the morning Renate organised the strike.'

Other sentences discarded from Kaan et al's original set contained either potential ambiguities or adjuncts that differed in semantic function between conjuncts.

In the original sentences, the objects, such as *het gazon* in (3a), are all definite expressions. Since we expected that non-generic objects would be more difficult to interpret in the proposed conditions *c* and *d*, we changed them to indefinite objects – where possible. As we can see in (6), an object that refers to exactly one of a set may cause an odd reading when it is elided in the right conjunct.

(6) a. Nina arranged the grill and Ruben hooked up the tap.
b. ?Nina arranged the grill and Ruben too.
c. Nina arranged a grill and Ruben too.

In (6b), it is hard to believe that one and the same grill is arranged twice, while in (6c) it is plausible that two people arranged two grills separately. The difference here is easily explained in terms of the definiteness of the NPs. Definite NPs in (6a) and (6b) refer to unique (some scholars use the term "familiar") entities in the context. Note, that the difficulty caused by uniqueness does not (immediately) arise with so-called weak definites such as *het gazon* in (3a) above.

Again on the basis of material used in Kaan et al. (2013), fillers were designed. (7a) is an example of a plausible filler and (7b) is an example of an implausible filler.

(7) a. Terwijl Gerda op de bank televisie keek, zat Sanne aan tafel te puzzelen.
While Gerda on the couch television watched sat Sanne at table to puzzle
'While Gerda watched TV on the couch, Sanne solved a crossword at the table.'

b. Nadat Esmee de post bij de villa bezorgde, keek de hond luid naar haar.
After Esmee the mail at the villa delivered looked the dog loudly at her
int: 'After Esmee delivered the mail at the villa, the dog looked at her loudly.'

While all test sentences started with the conjunction *omdat* 'because', fillers started with *omdat* 'because', *aangezien* 'since', *doordat* 'as a result of', *nadat* 'after', *voordat* 'before', or *terwijl* 'while'. Fillers differed in word length between 9 and 21 words. Thirty-six plausible fillers and 32 implausible fillers were constructed. A full list of the stimuli of this pretest can be found in Appendix B.

As discussed in section 3.3, complexity in ellipsis is subject to inconclusive evidence, Copy α and the cue-based mechanism predicting comparable results. However, in this design, it is not the form of the antecedent which changes but the complexity of to be recovered material, which is possible when using Gapping-like constructions. This allows us to compare different sizes of structure elisions within one sentence. In line with the suggestion of Poirier, Wolfinger, Spellman, and Shapiro (2010), we hypothesise that if more structure is elided, this might affect processing load.

Participants

Twenty participants participated and received €3 for their cooperation. Two participants did not obey the instructions: one took too much time to complete the experiment, the other appeared to have misunderstood the task. Two additional participants were invited as substitutes. The results below are based on twenty participants (four male; $M_{\text{Age}} = 24.45$, range 18-49).

Procedure

The items were divided over four lists using a Latin Square design. Each list contained only one member of each quadruplet and each participant rated only one list. The stimuli, which were interspersed with the 68 fillers described above (32 uninterpretable and 36 interpretable), were presented in an individually randomised order using the software PsychoPy (Peirce, 2007, 2009). Uninterpretable sentences had a well-formed structure but contained mismatching lexical items. Participants were asked to rate the sentences on a seven-point scale (see section 3.1 for a discussion on acceptability tests). They were encouraged to take into account the structure as well as the interpretability of the presented sentences. Also, they were asked to react as quickly as possible to obtain intuitive responses. Before the actual test, which contained 112 sentences, participants completed a practice session of 21 sentences. The experimental session took 25 minutes at the most.

5.2.3 Results

The mean ratings were calculated per quadruplet and per sentence. Quadruplets of test sentences of which one item had an average score below 4 were disregarded. Since the stimuli would be counterbalanced in the subsequent ERP experiment so that each participant only saw one sentence of a quadruplet, the number of quadruplets should be dividable by 4. Of the 38 remaining quadruplets an additional 2 quadruplets were removed on the basis of lowest scores per quadruplet and per sentence. After applying these criteria, thirty-six quadruplets remained for the following analysis. One implausible filler sentence was rated 5.20 on average. This filler was excluded along with the eight discarded quadruplets.

Condition	Mean	N	Standard Error
a	5.63	36	.11
b	5.49	36	.09
c	5.15	36	.11
d	5.17	36	.10
Total	5.36	144	.05

Table 5.1: Means of rating of test sentences per condition after correction.

In Table 5.1 the average ratings of the remaining test sentences are listed. The mean rating of plausible and implausible filler sentences was $M = 6.55$ [$SE = 0.07$] and $M = 2.19$ [$SE = 0.06$], respectively. The means of the test sentences were evaluated using a one-way ANOVA. Between four test conditions a main effect of CONDITION was found, [$F(3, 140) = 5.21, p = .002, \eta_p^2 = .100$]. A Bonferroni post hoc analysis of planned contrasts revealed that condition *a* differed marginally from condition *b* [$p = .069$], but it differed significantly from condition *c* [$p = .011$] and condition *d* [$p = .022$]. No other significant differences were apparent.

5.2.4 Discussion

The stimuli in this acceptability test consisted of plausible fillers, implausible fillers, and test sentences – the items of main interest. Relative to the control condition, the test conditions displayed a decline in ratings as more and more structure was elided. As expected, condition *a*, the control condition without ellipsis, was rated the highest while sentences with more elided structure were judged lower. Especially the inclusion of an object in the ellipsis (conditions *c*-*d*) had an effect on the mean ratings. Note though, that the steps between conditions *b, c, d* were not significant. Notably, numerically, the difference between the Gapping condition *c* and the subtype of Gapping (Stripping) condition *d* in which more structure was elided was almost equal.

The decreasing ratings relative to the control condition could be related to the "amount of repair" of structure as discussed in Chapter (1). In that sense, more elided structure may amount to a relative processing cost, while Stripping constructions (condition *d*) might be easier to repair than Gapping constructions. It will be interesting to see to what extent a processing cost affects comprehension of elliptical sentences and how this is reflected in terms of ERPs. In the ERP experiments in which a comprehension task will be included I will try to establish this.

One may ask why the test sentences were generally judged less acceptable than the plausible fillers. A tentative explanation could be that the test sentences consisted of three clauses instead of two as is the case in the fillers. Possibly, participants found sentences with more clausal content more difficult. During the debriefing of the experiment some of the participants indeed pointed to the issue of "too much information" in one sentence. Additionally, corpus research could be helpful to investigate to what extent the form of the test items differs from that of the filler items in terms of frequency. Stimuli with relatively more elided structure were rated relatively low. As mentioned above, this could be down to processing cost, but it could also be that such sentence forms are not frequently used. Note that low frequency items usually correlate with processing difficulty (see for example Levy, 2008).

5.2.5 Conclusion

The goal of this pretest was to check which of the quadruplets, that were designed on the basis of the first ERP experiment on Dutch Gapping, could be used in the planned ERP experiments reported in this thesis. By conducting a computer administered experiment in which the test sentences were presented together with plausible and implausible fillers, thirty-six of 44 quadruplets appeared to have adequate acceptability ratings. This means that these stimuli are considered as acceptable by native speakers of the language in terms of structure and interpretation. Compared to the control condition, a tendency of acceptability to decline as more structure is elided was observed. This could indicate that, when relatively more structure has to be recovered, processing load increases. Using the pretested stimuli in ERP experiments, I will try to shed light on the nature of processing mechanisms. Additionally, I will be able to compare acceptability judgement data from this pilot to comprehension data that will be collected and analysed in Chapter 6.

5.3 Norming stimuli II: acceptability of quantifiers

5.3.1 Purpose

In this norming study, proposed test sentences with semantic difficulty were tested for acceptability by native speakers. Items were included to compare the quantifiers *elke* "every" and *alle* "all" with the determiner *de* "the" in Gapping conditions and Stripping conditions. The latter modulation is tested in the ERP experiment reported in Chapter 7. In other items the additive marker *ook* "too" contrasts with the polarity marker *niet* "not". These items are included for follow-up experiments (not reported in this thesis).

5.3.2 Method

Participants

Forty native speakers of Dutch (10 male; $M_{\text{Age}} = 22.24$, range 19-31) participated and received €5 compensation.

Stimuli

On the basis of the original data set used by Kaan et al. (2013), ninety-five quintuplets as in (8) below were designed.

(8) a. Koen verving de kast in de woonkamer, en Judith de lamp in
Koen replaced the cabinet in the living.room and Judith the lamp in
de gang.
the hall
'Koen replaced the cabinet in the living room, and Judith the lamp in the hall.'

b. Koen verving elke kast in de woonkamer, en Judith de lamp
Koen replaced every cabinet in the living.room and Judith the lamp
in de gang.
in the hall
'Koen replaced the cabinet in the living room, and Judith the lamp in the hall.'

c. Koen verving de kast in de woonkamer, en Judith niet.
Koen replaced the cabinet in the living.room and Judith not
'Koen replaced the cabinet in the living room, and Judith did not.'

d. Koen verving de kast in de woonkamer, en Judith ook.
Koen replaced the cabinet in the living.room and Judith too
'Koen replaced the cabinet in the living room, and Judith too.'

e. Koen verving elke kast in de woonkamer, en Judith ook.
Koen replaced every cabinet in the living.room and Judith not
'Koen replaced every cabinet in the living room, and Judith too.'

As I have explained in Chapter 2.4.2, quantifying expressions may be a burden on mechanisms of movement and/or copying since additional structural information has to be analysed. Therefore, I created stimuli to test the difference between quantified phrases and phrases containing a definite article. Condition *a* is the same as the plausible Gapping condition that was used in the replication of Kaan et al. 2013 reported earlier. This condition contrasts with condition *b* in which the determiner of the object in the left conjunct is replaced by a quantifier. In condition *c*, the negative polarity marker at which the ellipsis is resolved can be compared to the (positive) additive marker in condition *d*. In turn, condition *d* can be contrasted with condition *e* to estimate the difference between a determiner and a quantifier in Stripping constructions. The latter comparison will be further explored in Chapter 7 which reports an ERP experiment that was designed to focus on the semantic aspect of retrieval and integration processes.

Procedure

The items were counterbalanced over five lists. Each list contained only one member of each quintuplet and each participant rated only one list. The stimuli, which were interspersed with an additional 93 fillers of which 22 uninterpretable, were presented in an individually randomised order using the software PsychoPy (Peirce, 2007, 2009). Uninterpretable sentences had a well-formed structure but contained mismatching lexical items (similar to the uninterpretable items used in the pretest described above). Participants were instructed to take into account the structure as well as the interpretability of the presented sentences, and to rate the sentences on a seven-point scale. To obtain intuitive responses, they were asked to react as quickly as possible. Before the actual test, which contained 188 sentences, participants completed a practice session of 21 sentences. The session lasted 30 minutes on average.

5.3.3 Results

Due to a scripting error, three conditions of one stimulus set were wrongly coded and presented as the same condition. Therefore, the analysis is based on the remaining 94 stimuli sets. The mean ratings were calculated per sentence. In Table 5.2, the means and standard errors are listed for the five test conditions with mean scores higher than 4.

Condition	Mean	N	Standard Error
a	5.85	92	.07
b	5.16	75	.07
c	5.23	80	.06
d	5.40	86	.07
e	4.99	70	.08
Total	5.36	403	.04

Table 5.2: Means of rating of test sentences per condition after correction.

The mean ratings of plausible and implausible filler sentences were $M = 6.55$ [$SE = 0.07$] and $M = 2.80$ [$SE = 0.06$], respectively. Table 5.2 shows that low mean scores coincide with a relatively high exclusion rate. In general, the items containing quantifiers were judged least acceptable. Since conditions *d* and *e* are to be tested in the ERP experiment reported in Chapter 7, these items were analysed in more detail. From the data set, 42 pairs of conditions *d* and *e* were chosen such that they matched in terms of their mean ratings. Items within such a pair maximally differ in 1.25 average acceptability score points. The range of average scores among chosen items was 4.38-6.50; means of condition *d* [$M = 5.46$, $SE = 0.10$] and condition *e* [$M = 5.32$, $SE = 0.08$] did not differ significantly [$t(41) = 1.41$, $p = .166$, $d = .218$].

5.3.4 Discussion

Since the sentence structures in this norming study more closely resemble the original stimuli tested by Kaan et al. (2013) than the stimuli in the first norming study, it was easier to construct a larger set of stimuli. As a consequence, a set consisting of conditions *d-e* could be chosen in which the means differ minimally. Note that therefore we have the luxury of controlling the effect of acceptability in subsequent experiments using these stimuli, but this is not possible for the stimuli set derived from the first norming study. At least numerically, the sentences with quantifiers were rated lower than the other conditions, followed by condition *c*, which contained negation at the ellipsis site. In this sense, semantic difficulty seems to correlate with lower ratings, that is, acceptability may decrease as a function of semantic complexity.

As was the case in the first norming study, the elliptical sentences were rated lower than the plausible fillers. It was proposed in the first norming study that this may be down to the inclusion of three clauses in the sentence structure. Since the elliptical sentences in the current set do not have this property, it may in fact be the case that ellipsis is less acceptable than fully-fledged sentences in general. It should be noted though that "acceptability" is not only a measure of grammaticality but it is also dependent on the relative difficulty of interpretation and therefore likely related to a relative processing cost that may resemble the resolution process. In the subsequent ERP experiments, this will be investigated in more detail.

5.3.5 Conclusion

A norming study was carried out to ascertain the acceptability of stimulus sentences containing Gapping and Stripping constructions which differed in terms of semantic complexity. From the pool of tested sentences a set has been selected for use in the ERP experiment described in Chapter 7, where semantic complexity is investigated. In contrast to the result of the first norming study, a set could be compiled in which the means of acceptability differed only minimally. Consequently, the factor ACCEPTABILITY need not be considered as factor in the ERP experiment on semantic complexity.

Additional stimulus sets that have been tested in this section may be used in future experiments – for example, as a follow-up of the current thesis (e.g. a comparison between the additive markers *ook* and *niet* to investigate negated elisions). In the remaining chapters, however, we will be concerned with the four ERP experiments that have been conducted.

CHAPTER 6

ERP experiments I & II: Structural complexity

This chapter reports two ERP experiments that focus on the effects of the amount of elided structure. I am thankful to Naomi Nota, Olga Kepinska and Ferdi van de Kamp for assistance during some parts of the data collection.

6.1 Modulation of structure in the right conjunct

6.1.1 Introduction

The aim of the experiment reported in this section was to find out the effect of the amount of structure that is elided in gapping-like constructions. In a word-by-word reading task, EEG was recorded while at the same time participants' comprehension was measured. Condition by condition a phrase is cut off. At a fixed measure point across conditions, ERPs were recorded to analyse the effect of increasing amounts of structure. A memory task was included to control for individual variation in relation to memory retrieval (c.f. Kaan et al., 2013). The stimuli used in this experiment were rated by other participants for their acceptability in a computer-administered judgement task (see Section 5.2.2 for the rationale behind the experimental sentences).

At the critical measure point, a copy account predicts a processing cost relatively early, that is, just after encountering a gap. Possibly this would be reflected by a LAN or ELAN component. The subsequent integration of retrieved structure should be relatively easy. A cue-based account, however, predicts the reverse: less retrieval cost and possible ERP effects relatively late in the time course; for example, a P600 reflecting an integration cost.

6.1.2 Methods

Test materials

As explained in chapter 5.2.2, thirty-six quadruplets as in (1) were chosen. For a complete list of test sentences and average acceptability scores see Appendix A. A Latin square design was applied to counterbalance the stimuli so that each participant only saw one sentence of a quadruplet. Along with the 36 test sentences, 72 filler sentences, half of which containing proper names, were added. The test sentences are shown in (1) (repeated from chapter 5.2.2). The critical word *waren* is in bold: ERPs were measured in relation to the presentation of this word. The colour of *waren* corresponds with the type of condition.

(1) a. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin de paden harkte, **waren** de buurtgenoten
in the back.garden the paths raked were the neighbours
vrolijk.
happy
'Because Hilde maintained the lawn in the front garden
and Ralph raked the paths in the back garden,
the neighbours were happy.' (***No Gapping***)

b. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin de paden, **waren** de buurtgenoten vrolijk.
in the back.garden the paths were the neighbours happy
'Because Hilde maintained the lawn in the front garden
and Ralph the paths in the back garden,
the neighbours were happy.' (***Verb Gapping***)

c. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin, **waren** de buurtgenoten vrolijk.
in the back.garden were the neighbours happy
'Because Hilde maintained the lawn in the front garden
and Ralph in the back garden,
the neighbours were happy.' (***Verb-Object Gapping***)

d. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
ook, **waren** de buurtgenoten vrolijk.
too were the neighbours happy
'Because Hilde maintained the lawn in the front garden
and Ralph too,
the neighbours were happy.' (***Stripping***)

Note that I use a comma to separate clauses as is common in Dutch, especially between two finite verbs. In general, a comma is beneficial for the reader as is suggested by Yang (2010) (see also https://onzetaal.nl/taaladvies/komma-voor-en). I was not interested in any ambiguity or mismatch effects at the point that the main verb of the sentence is processed in Gapping sentences, therefore a second comma is helpful because it would prohibit such possible effects. Also, the finite verb of the main clause consistently disagrees with the subject of the second conjunct in terms of number. The subject of the main clause *de buurtgenoten* is plural, in contrast to the singular subject in the preceding clause *Ralph*. Following Kaan et al. (2013) an additional comma was placed before *en* to ensure that the noun of the right conjunct would not be erroneously interpreted as the second object of the preceding verb.

Participants

Twenty-nine native speakers of Dutch with normal or corrected-to-normal vision took part in this study and were paid €15. Six of them were disregarded from the analysis because they were left-handed (1 participant), because of technical failure (1 participant) or because of too many artefacts according to the (stringent) criteria described below (4 participants). Of the remaining 23 participants, 13 were female and the mean age was 23.26 (age range 19-37). The experiment followed the Ethics Committee regulations of the Humanities Faculty of Leiden University, which approved its implementation. Participants gave informed consent before the study.

Procedure

Participants were comfortably seated in a dimly lit sound-proof room at a distance of approximately 80 cm of a 17 inch CRT monitor. One-hundred-and-eight test sentences were presented in a random order using the presentation software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The sentences were presented word by word in Verdana font (18pt). Each word was presented for 300 ms with a 300 ms fixation cross interval. Presentation of a trial started with a fixation cross for 1,000 ms. Every sentence was followed by a content question to encourage comprehension. A blank screen with a duration of 1,500 ms intervened between the last word of a trial and the presentation of the comprehension question. For half the participants the left response button referred to "YES", for the other half the left button referred to "NO". Participants were given a break after 12 sentences and could proceed at their own pace. The comprehension questions referred to different parts of the sentences equally. Before the actual test, the participants were able to get used to the task with four practice sentences.

The experiment was concluded with a working memory test based on a task described in chapters 4.2.3 and 5.1.1.

The experiment took about 1.5 hours per participant, including EEG set-up.

Apparatus and electrophysiological recording

A description of the recording set-up can be found in chapter 5.1.1.

Data analysis

The EEG data were preprocessed using Brain Vision Analyzer Version 2.0. (Brain Products, Munich, Germany). EOG artefacts were corrected using the Gratton et al. (1983) algorithm. Remaining artifacts were rejected on the basis of the following criteria: trials below -200 μV, above 200 μV, or including a voltage step of 20 μV or more within 200 ms. A low cutoff filter of 0.05 Hz, 24bB/oct and a high cutoff filter of 30 Hz, 24 dB/oct were applied. Epochs of 1,000 ms were computed with a 200 ms pre-stimulus baseline. ERP grand averages were time-locked to the onset of the target word *waren*. As a result, 6.64 % of the trials were excluded from the analysis: of the nine trials presented per condition an average of 8.4 ($SE = 0.15$) per participant were used.

Both the behavioural data and the electrophysiological data were analysed by means of a repeated measures procedure in **R** using the `ez` package (Lawrence, 2011). To correct for possible sphericity violations in either analysis the Greenhouse-Geisser correction applies. Using the package `lmPerm` (Wheeler, 2010), time windows of interest were empirically determined by permutation tests that were run per sample at each electrode with the independent factor CONDITION from 0 to 800 ms after target onset.[1] This procedure is akin to running multiple t-tests (or ANOVAs) for consecutive, averaged 20 ms time windows (see for example Timmer & Schiller, 2012), however, the advantage of running a permutation test per sample is that the onset and offset of time windows can be determined with even more precision. Also, like non-parametric tests, permutation tests are robust for relatively small samples (Legendre & Legendre, 1998:20). For this thesis, I determined that consecutive significant samples with a duration of at least 50 ms occurring in at least one electrode were taken into consideration for follow-up analyses. These analyses were conducted separately for midline sites (Fz, Cz, Pz) and for the lateral electrode regions: left/right frontal (AF3/4, F7/8, F3/4), left/right central (FC1/2, C3/4, CP1/2), left/right parietal (P7/8, P3/4, PO3/4). In the current experiment a repeated measures ANOVA was planned using within-subjects factors CONDITION (four levels: No Gapping, Verb Gapping, Verb-Object Gapping, Stripping), ANTERIORITY (3 levels: frontal, central, posterior), and, for analyses involving lateral sites, HEMISPHERE (2 levels: left, right).

6.1.3 Behavioural results

The accuracy on the comprehension questions was on average 92.75% ($SE = 1.04\%$). Per condition the accuracy scores were as follows: $M_{\text{No Gapping}} = 88.41\%$, $M_{\text{Verb Gapping}} = 94.69\%$, $M_{\text{Verb-Object Gapping}} = 91.79\%$, and $M_{\text{Stripping}} = 96.14\%$. The accuracy scores differed significantly between conditions as a repeated measures on the ratios showed [$F(3, 66) = 3.50$, $p = .039$, $\eta^2_G = .089$]. A post hoc comparison with Bonferroni correction was carried out to check which conditions differed specifically. Only between conditions No Gapping (***a***) and Stripping (***d***) a notable difference was found, but this did not reach significance [$p = .088$].

Due to a scripting error, the first session of the working memory task (random counting) consisted of four trials instead of five. The accuracy ratio of

[1] I am very thankful to Cesko Voeten who provided a script to make this work, since it made the procedure of combined data analysis in Brain Vision Analyzer and **R** workable.

the three test sessions was on average 65.83% ($SE = 2.65\%$). Per condition the scores were: $M_{\text{Random Counting}} = 66.30\%$, $M_{\text{Auditory Presentation}} = 60.00\%$, $M_{\text{Visual Presentation}} = 71.30\%$. A repeated measures ANOVA by subjects with CONDITION as independent factor and ACCURACY OF NUMBER RECALL as dependent variable showed that the ratios did not differ between conditions [$F(2, 44) = 1.55, p = .223, \eta^2_G = .045$].

The scores of the comprehension task of the ERP experiment were compared with the scores of the working memory task. A small, non-significant correlation was found between the variables ACCURACY OF SENTENCE COMPREHENSION and ACCURACY OF NUMBER RECALL [$r = .249, p = .251$].

6.1.4 Electrophysiological results

Onset of *waren*

In Figure 6.1, the mean amplitudes of all conditions are depicted as measured from the onset of *waren*. Particularly conditions Verb-Object Gapping (***c***) and Stripping (***d***), show a large and sustained negativity as compared to the No Gapping condition (***a***). However, deviations appear to already begin before onset, yielding an erratic baseline. Analysis of effects after such a distorted baseline are unreliable, since they may be attenuated by effects earlier on.[2]

6.1.5 Towards an alternative measure point

The accuracy of the comprehension questions showed that the participants understood the stimuli almost perfectly in all conditions; nonetheless the fewest mistakes were made in condition ***d***, which may not be surprising considering that this condition contained the least amount of "new information" in the right conjunct. Sentence comprehension barely correlated with the working memory scores. It is notable that, while the acceptability study reported in 5.2 showed that condition ***d*** was rated relatively the lowest (together with condition ***c***) in terms of acceptability, it was the best understood according to the comprehension scores of the current experiment. Possibly, the issue of the least amount of new information again plays a role. The comprehension task requires that all of the information of the foregoing sentence remains accessible in memory, which is easier when there is less (new) information. In addition, it could be that the sub-type of Gapping, namely Stripping (as in condition ***d***), is more frequently used than Gapping (as in conditions ***b***-***c***) in Dutch.

A limitation of this study concerns my decision to measure ERPs at the main verb following the subordinate clause, *waren*. This point was chosen because an earlier measure point, immediately before *waren*, would face the problem of different and hence incomparable phrases across four conditions.

[2] I am thankful to the people that attended my poster at the CNS meeting in Boston in 2014 and who politely pointed out my classic rookie mistake – overlooking the erratic baseline.

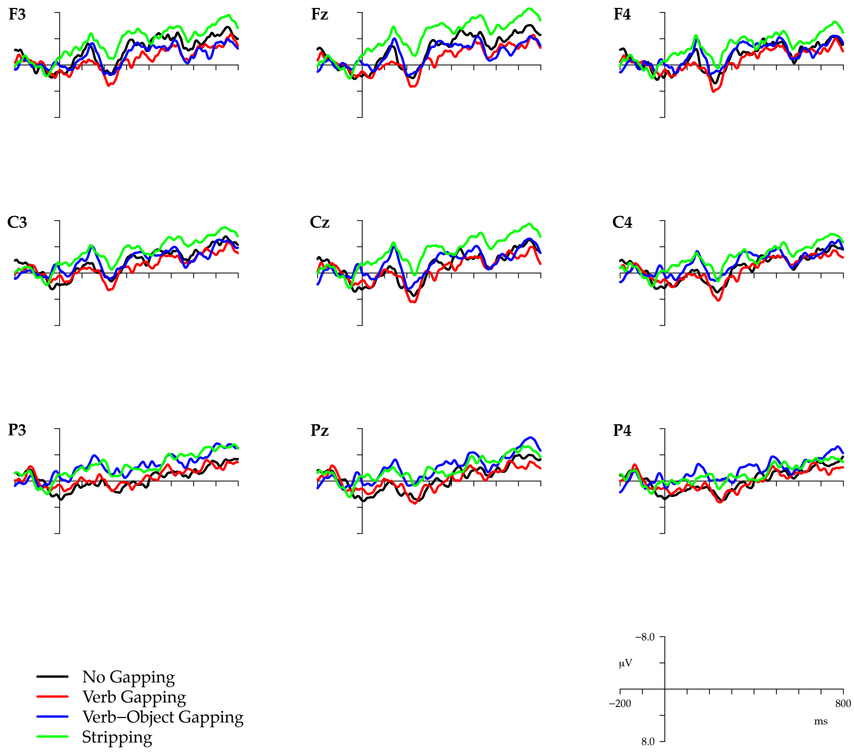


Figure 6.1: Grand averages of all conditions at the onset of *waren* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 109.

However, my planned measure point appeared to be too late since a proper baseline could not be determined, as explained above. It may be the case that once the elision (or *ook*, for that matter) was being processed, the search for the antecedent started immediately. Also, the comma that was presented with the word before the main verb is a sign for the reader that the resolution process may start, as it marks a clause boundary. This process could have caused a so-called "spill-over effect", as has been noted by Steinhauer and Drury (2012): an effect elicited before the onset of *waren* prevents us from drawing conclusions about any possible component elicited at the onset of *waren*. Steinhauer and Drury (2012) used the spill-over effect to explain unjustified ELAN effects and it may also be applicable for the current study. It seems, therefore, that we tapped into the recovery process at least 300 ms late. Such a delay could even mean missing the recovery process in a condition such as ***b***, in which the ellipsis was small and could potentially have been resolved before *waren*.

A solution to overcome spill-over effects is to use a different baseline, for example, the proper name in the second conjunct. However, this would lead to analysing large time windows, especially in conditions ***a***-***b***. With the current data, this would in turn lead to excluding a considerable amount of trials because of an increased number of artefacts. Considering the already small number of trials in the current design, calculating an alternative baseline was not an option. Additionally, the comma marks a clause boundary as well a prosodic boundary. It is conceivable that the comma induced the ellipsis resolution process. Therefore, it would be even more appropriate for us to find out whether differences between conditions are apparent on words that appeared just before *waren* and that can be compared with appropriate counterparts. In the next section, I will explore a post hoc analysis of such words.

Post hoc analysis of comma effects

It is possible to compare nouns with a comma in conditions ***b*** and ***c*** – amounting to Gapping constructions – with their counterparts which are not accompanied by a comma in the No Gapping condition ***a***. Two different nouns can be taken into consideration for comparison. The first is the object noun. Below I repeat the example of the stimulus set. In (2) it is shown that between conditions ***a*** and ***b*** the object nouns (in bold) can be compared. I hypothesise that an ERP effect can be measured from the onset of ***paden*** in condition ***b*** as compared to condition ***a***, showing the effect of the resolution process of the elided verb of the second conjunct.

(2) a. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin de **paden** harkte, waren de buurtgenoten
in the back.garden the paths raked, were the neighbours
vrolijk.
happy
'Because Hilde maintained the lawn in the front garden
and Ralph raked the paths in the back garden,
the neighbours were happy.' (***No Gapping***)

b. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de achtertuin de **paden**, waren de buurtgenoten vrolijk.
in the back.garden the paths, were the neighbours happy
'Because Hilde maintained the lawn in the front garden
and Ralph the paths in the back garden,
the neighbours were happy.' (***Verb Gapping***)

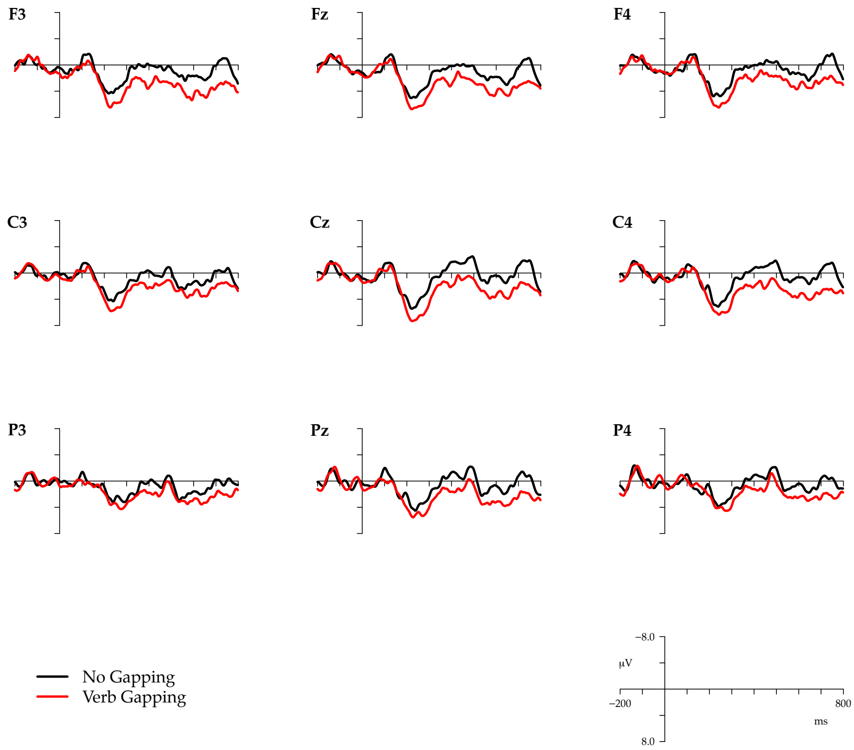


Figure 6.2: Grand averages of conditions No Gapping (***a***) and Verb Gapping (***b***) at onset (y-axis) object noun (*paden*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 114.

Plots of the grand averages were evaluated together with a permutation test per sample per electrode. Time windows of interest were empirically determined between 220-520 ms and 570-780 ms. As a check, a repeated measures ANOVA was run for the intermediate time window of 520-570 ms. No effects could be found for the factor CONDITION on midline sites [$F(1, 22) = 1.83$, $p = .189$, $\eta^2_G = .027$] and on lateral sites [$F(1, 22) = 1.18$, $p = .288$, $\eta^2_G = .017$]).

Object noun: effects in the 220-520 ms time window

On midline electrodes, the factor CONDITION reached significance [$F(1, 22) = 5.37$, $p = .030$, $\eta^2_G = .070$]. Additionally, a marginal effect of ANTERIORITY was found [$F(2, 44) = 3.69$, $p = .054$, $\eta^2_G = .013$]. A pairwise comparison showed a significant difference between frontal and central electrodes [$p = .007$, $M_{\text{Frontal}} = 1.83$ ($SE = 0.56$), $M_{\text{Central}} = 0.85$ ($SE = 0.58$)]. No interaction between CONDITION and ANTERIORITY was found. On lateral electrodes, only an effect of CONDITION could be established [$F(1, 22) = 4.54$, $p = .044$, $\eta^2_G = .054$].

Object noun: effects in the 570-780 ms time window

The effect of CONDITION was similar between midline electrodes [$F(1, 22) = 4.87$, $p = .038$, $\eta^2_G = .077$] and lateral sites [$F(1, 22) = 4.88$, $p = .038$, $\eta^2_G = .066$]. No other effects or interactions were found.

Adjunct noun: effects in the 300-430 ms time window

The second noun that can be considered is the noun in the adjunct phrase *in de achtertuin*. (3) shows the relevant conditions. As compared to condition ***a***, it is predicted that in condition ***c*** an effect can be measured at ***achtertuin***, representing the recovery of the VP *het gazon onderhield*.

(3) a. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de **achtertuin** de paden harkte, waren de buurtgenoten
in the back.garden the paths raked, were the neighbours
vrolijk.
happy
'Because Hilde maintained the lawn in the front garden
and Ralph raked the paths in the back garden,
the neighbours were happy.' (***No Gapping***)

c. Omdat Hilde in de voortuin het gazon onderhield, en Ralph
Because Hilde in the front.garden the lawn maintained and Ralph
in de **achtertuin**, waren de buurtgenoten vrolijk.
in the back.garden, were the neighbours happy
'Because Hilde maintained the lawn in the front garden
and Ralph in the back garden,
the neighbours were happy.' (***Verb-Object Gapping***)

Plots of the grand averages were evaluated together with a permutation test per sample and per electrode. A time window of interest was determined between 300-430 ms.

On midline electrodes, the factor CONDITION reached significance [$F(1, 22) = 4.83$, $p = .039$, $\eta^2_G = .048$] as well as on lateral sites [$F(1, 22) = 6.69$, $p = .017$, $\eta^2_G = .059$]. No other effects or interactions could be found.

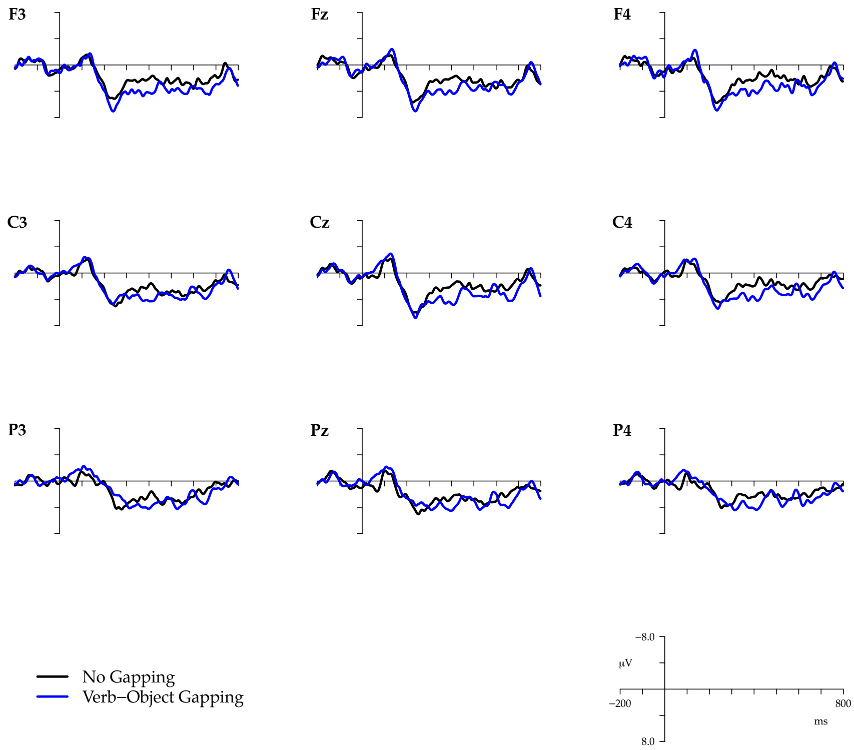


Figure 6.3: Grand averages of conditions No Gapping (***a***) and Verb-Object Gapping (***c***) at onset (y-axis) adjunct noun (*achtertuin*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 116.

6.1.6 Discussion and conclusion

In this experiment, I attempted to tap into the resolution process of Gapping-like constructions while manipulating the amount of structure to be recovered. As it turned out, my planned measure point was not an appropriate critical time point. In a post hoc analysis of time points earlier in the experimental sentences, it appears that indeed effects of CONDITION could be established. Effects are apparent before the intended measure point *waren*.

Rather counter-intuitively, the recovery of the verb *onderhield* in condition ***b*** yielded a larger effect (lasting for 510 ms in total) compared to the recovery of the VP *het gazon onderhield* in condition ***c*** (lasting for 130 ms in total). Although it is possible that the sustained positivity in condition ***b*** is caused by individual variation, I will argue that it is composed of two consecutive positive deflections.

Regardless of whichever account one wants to relate the results, it was expected that the elision of a larger structure would result in a larger effect overall. With the current results we could provisionally conclude that the recovery of a VP is more effortful than to connect a fully elided VP with a remnant adjunct phrase. Thus, in the post hoc analysis it appears that it depends on the type of elided structure (quality) rather than the amount of structure (quantity). To recover a VP a predicate argument relation between elided verb and remnant object needs to be established. It could be that this relation is readily available in case a full VP is retrieved. The integration of a full VP with a remnant adjunct phrase is predicted to be relatively easy by structural accounts (including Copy α) which assume that an adjunct attaches directly to a VP node. However, it is assumed to encounter relatively more problems in searching for the exact structure in memory. A cue-based approach, that may translate different information types as cues, predicts the reverse: relative ease of retrieval, but higher integration cost as different information types need to be matched. Therefore, the two proposed mechanisms may both account for the relative ease of processing an elided VP and adjunct phrase. It is still unclear how these mechanisms may help to explain why the recovery of a VP is relatively more costly.

In both manipulated conditions, the effect of CONDITION amounted to positive deflections starting around 300 ms. Further, it seems that two consecutively positive components can be related to a relatively effortful resolution process. This contrasts with the negativity as found on the determiner (that was followed by a P600 on the subsequent noun) in the replicated study in Chapter 5.1. The first positive component found in the current experiment could be in part connected to a closure positive shift (CPS). It has been argued that the appearance of a comma triggers subvocal prosodic sentence phrasing – marking a prosodic boundary (Drury, Baum, Valeriote, & Steinhauer, 2016; Steinhauer & Friederici, 2001). Yet, the latency of the CPS may differ as a function of task or language used in the experiment (see for example Peter, McArthur, & Crain, 2014). In that sense, a negativity as found in the replication study could be outweighed by a subvocal prosodic effect, yielding a different polarity. Note that the negativity was found on a function word – a determiner – which is usually not assigned meaningful prosodic content, contrasting with remnants that survive by virtue of focus assignment. However, if it were only a reflection of subvocal prosody, we should have found a similar latency and duration of the component in all conditions. As this is not the case, I suggest that the early positivity is sustained by both 'acoustic' and linguistic cues (see Peter et al., 2014 for a discussion that both acoustic and linguistic cues are involved in the CPS in 'normal' sentences). The early positivity marks the start of the resolution process to retrieve missing information, to integrate it with the remnant structure and to arrive at an interpretation. I further suggest that the relative difficulty of integration of retrieved material is then reflected by the secondary positivity which I relate to a P600.

Besides the CPS, the early positivity could be a family member of a more domain-general P300 which has been related to surprisal effects (in so-called "oddball tasks") and context updating (see for example Donchin, 1981; Donchin & Coles, 1988). In particular, in an overview and discussion of the P300, (Polich, 2007:2130) states that the amplitude and latency of a P300 varies as a function of task demands to the extent that "for tasks that require greater amounts of attentional resources, [the] P300 amplitude is smaller and peak latency is longer". While a context updating account may be in line with a linguistic resolution process, it should be noted that a P300 may be simply reflecting attentional, evaluative, or memory mechanisms (see for a review Nieuwenhuis, Aston-Jones, & Cohen, 2005). Still, in the experimental linguistic literature, Gouvea et al. (2010) reported that an anterior positivity between 300-500 ms was elicited by a condition during which a *wh*-dependency was resolved, marking the onset of a retrieval process. These authors also suggest that "manipulations that impact the number and type of syntactic relations that are attempted should change the amplitude and/or duration of the P600" (2010:174), which may be in line with the secondary positivity as established on the object noun.

However, a caveat is in order. At the critical nouns, the clauses are "wrapped up" up to that point. This means that retrieved information is integrated with the remnants. In condition ***b*** it can be seen that there is more "new" information than in condition ***c***. This extra information may have caused a spill-over effect, increasing the demands on the integration process. While this cannot be explained by a copy account, it can be accounted for by means of cues, as new information – say, every contrasting remnant – may interfere with old information. It has been shown that a cue-based account is apt to provide an explanation for interference effects (see for example Martin & McElree, 2008). As a consequence of the difference in information load, it is difficult to compare the effect on the object noun with the effect on the adjunct noun. In that sense, the effect does depend on a spill-over effect rather than on the nature of the elided structure. In the following experiment, I control for a possible spill-over effect.

6.2 Modulation of structure in the left conjunct

6.2.1 Introduction

Still being concerned with the effect of elision of different structure size and in order to control for spill-over effect in the second conjunct, I designed a follow-up ERP experiment. In this experiment, I utilise Stripping since it will help to control for the amount of new information in the second conjunct. In Stripping, the structure to be elided can be manipulated in the left conjunct while the size of the remaining structure is constant in the right conjunct. Let us have a look at the stimuli to further explain the logic and hypotheses for the current design.

6.2.2 Methods

Test materials

On the basis on the Stripping condition of the previous experiment a stimulus set of 36 quadruplets was compiled as in (4):

(4) a. Omdat Koen een kast verving, en Judith **ook**, waren de bewoners perplex.
Since Koen a cabinet replaced and Judith too were the inhabitants perplexed
'Since Koen replaced a cabinet and Judith too, the inhabitants were perplexed.' (***VP Stripping***)

b. Omdat Koen een enorme kast verving, en Judith **ook**, waren de bewoners perplex.
Since Koen a huge cabinet replaced and Judith too were the inhabitants perplexed
'Since Koen replaced a huge cabinet, and Judith too, the inhabitants were perplexed.' (***VP-Adjective Stripping***)

c. Omdat Koen in de woonkamer een kast verving, en Judith **ook**, waren de bewoners perplex.
Since Koen in the living.room a cabinet replaced and Judith too were the inhabitants perplexed
'Since Koen replaced a cabinet in the living room, and Judith too, the inhabitants were perplexed.' (***VP-Adjunct Stripping***)

d. Omdat Koen in de woonkamer een enorme kast verving, en Judith **ook**, waren de bewoners perplex.
Since Koen in the living.room a huge cabinet replaced and Judith too were the inhabitants perplexed
'Since Koen replaced a huge cabinet in the living room, and Judith too, the inhabitants were perplexed.' (***VP-Adjective-Adjunct Stripping***)

As can be observed, a baseline was constructed in condition ***a*** in which the

VP *een kast verving* is stripped in the second conjunct and is resolved at the word *ook*, which will be the critical measure point. In the first non-baseline condition (***b***) the adjective *enorme* is added to the object noun of the VP, to expand the VP. Condition ***c*** in this set equals the Stripping condition of the experiment reported in section 6.1 and contains the VP of condition ***a*** with the addition of the adjunct *in de woonkamer*. In the last manipulation in the current set, the baseline VP has the addition of both the adjective and adjunct. Clearly, the second conjunct has the same length in all conditions and consequently, contains the same amount of new information, as there is only one contrasting element (*Judith*). This was done to overcome the suggested spill-over effects during wrap-up and integration at the end of the clause. In the current design, an increase of information load is constrained to the elided structures.

Since the measure point will be at a word that will be presented with a comma, I predict – on the basis of the findings in the previous experiment – a positive component which reflects the effect of prosody and the start of the recovery process. The amplitude and duration of the positivity is further expected to be modulated by the amount of elided structure – the least in condition ***b*** and the most in condition ***d***. I further hope to corroborate previous effects in terms of two consecutive positive components. The first component, around 300 ms after onset of *ook*, would be understood as (partly) related to processes of retrieval. A secondary positive component occurring around 600 ms after onset will be considered as a measure of integration cost. I hypothesise that, if retrieval mainly depends on searching for and finding syntactic structure, a syntax-related ERP should be found. This is predicted by Copy α. Alternatively, retrieval cost that is reflected by modulations of the first positive component may be linked to several information types – not exclusively syntax-related. Possible modulations of integration cost as reflected by the second component may be connected to the cue-based retrieval account, since a copy account predicts a relatively easy integration.

The stimuli were interspersed with 72 fillers – the same as in the previous experiment – and were again counterbalanced using a Latin Square design.

Participants

Thirty-three native Dutch participants with normal or corrected-to-normal vision took part in this study and were paid €15. Two of them were disregarded from the analysis because of technical failure and three participants were taken out because of too many artefacts. Four participants were excluded because they were left-handed. Of the remaining 24 participants that were taken into consideration for further analysis, 8 were male and the mean age was 22.30 (range 19-33). The experiment followed the Ethics Committee regulations of the Humanities Faculty of Leiden University, which approved its implementation. Participants gave informed consent before the study started.

Procedure

Participants were comfortably seated in a dimly lit sound-proof room at a distance of approximately 90 cm from a 19 inch LCD monitor. One-hundred-and-eight test sentences were randomly presented word by word in Verdana font (36pt) for 300 ms per word with a 300 ms blank screen interval using the presentation software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Presentation of a trial started with a fixation cross for 1,000 ms. Every sentence was followed by a content question to ensure participants' attention. A blank screen with a duration of 1,500 ms intervened between the last word of a trial and the presentation of the comprehension question. For half the participants the left response button referred to "YES", for the other half the left button referred to "NO". Participants were given a break after every 12 sentences and could proceed at their own pace. The comprehension questions referred to different parts of the sentences equally. Before the test stimuli, the participants were able to get used to the task with four practice sentences.

The experiment was concluded with a working memory test as used in the previous experiment in this chapter.

In total, the experiment took about 1.5 hours per participant, including set-up.

Apparatus and electrophysiological recording

A description of the recording set-up can be found in chapter 5.1.1.

Data analysis

The EEG data were preprocessed using the same criteria as the previous analysis. Epochs of 1,000 ms were computed with a 200 ms pre-stimulus baseline and ERP grand averages were time-locked to the onset of the target word *ook*. 10.53% of the trials were excluded from the analysis: of the presented nine trials per condition an average of 8.05 ($SE = 0.13$) were used.

The repeated measures ANOVA on the ERPs had as independent factor CONDITION consisting of four levels (A, B, C, D) of which the first is understood as baseline. As in the other experiments, three levels of the factor ANTERIORITY (frontal, central, posterior) and two levels of the factor HEMISHERE (left, right) were taken into consideration.

6.2.3 Behavioural results

On average, the accuracy on the comprehension questions was 92.25% ($SE = 1.57\%$). Per condition the accuracy scores were: $M_{\text{VP Stripping}} = 92.60\%$, $M_{\text{VP-Adjective Stripping}} = 93.52\%$, $M_{\text{VP-Adjunct Stripping}} = 90.74\%$, and $M_{\text{VP-Adjective-Adjunct Stripping}} = 92.13\%$. The accuracy scores did not differ significantly between conditions as shown by a repeated measures ANOVA on the scores [$F(3, 69) = 0.42$, $p = .660$, $\eta^2_G = .004$].

The accuracy of the three test sessions of the working memory task was on average 68.61% ($SE = 2.45\%$). Per condition the scores were: $M_{\text{Random Counting}} = 62.50\%$, $M_{\text{Auditory Presentation}} = 59.17\%$, $M_{\text{Visual Presentation}} = 84.17\%$. A repeated measures ANOVA by subjects with CONDITION as independent factor and ACCURACY OF NUMBER RECALL as dependent variable showed that the scores differed between conditions [$F(2, 46) = 9.44, p < .001, \eta^2_G = .202$]. A post-hoc comparison with Bonferroni correction showed that the visual condition differed significantly from the random condition [$p = .014$] and auditory condition [$p < .001$].

Again, the scores of the comprehension task of the ERP experiment were compared with the scores of the working memory task. A small, non-significant correlation was found between the variables ACCURACY OF SENTENCE COMPREHENSION and ACCURACY OF NUMBER RECALL [$r = .213$, $p = .318$].

6.2.4 Electrophysiological results

In Figures 6.4 (on page 125), 6.5 (on page 126), 6.6 (on page 127), the grand averages of conditions ***b***, ***c***, ***d***, respectively are depicted relative to the baseline (condition ***a***). After evaluation of the graphs and a permutation test on all electrodes per sample with independent factor CONDITION, two time windows of interest were determined: 300-480 ms and 530-790 ms. In order to check the effects of the factor CONDITION in the intermediate time window 480-530 ms, a repeated measures ANOVA was run. No effects could be found on midline sites [$F(1, 23) = 0.87, p = .462, \eta^2_G = .017$] and on lateral sites [$F(1, 23) = 2.03, p = .120, \eta^2_G = .032$]).

Effects in the 300-480 ms time window

On midline sites, the factor CONDITION reached marginal significance [$F(3, 69) = 2.41, p = .074, \eta^2_G = .040$], while also a marginally significant interaction of CONDITION by ANTERIORITY was found [$F(6, 138) = 2.42$, $p = .059, \eta^2_G = .016$]. For convenience, the means of voltages per condition grouped by anteriority are shown in Figure 6.7 (on page 128).

On lateral sites, the factor CONDITION reached significance [$F(3, 69) = 4.82, p = .004, \eta^2_G = .050$]. A pairwise comparison with Bonferroni correction showed that both conditions ***a*** and ***b*** differed significantly from condition ***c*** and condition ***d***. Table 6.1 (on page 124) summarises these comparisons together with the values for the means, standard errors and *p*-values. While no interactions could be established, a marginal effect of HEMISPHERE was noticed [$F(3, 69) = 3.89, p = .061, \eta^2_G = .009$], as the left-localised electrodes were slightly more positive [$M_{\text{Left}} = 1.39$ ($SE = 0.20$), $M_{\text{Right}} = 0.81$ ($SE = 0.21$)].

Effects in the 530-790 ms time window

While the effect of CONDITION was marginally significant on midline electrodes [$F(3, 69) = 2.73$, $p = .051$, $\eta^2_G = .053$], a significant effect of ANTERIORITY was found [$F(2, 46) = 3.97$, $p = .036$, $\eta^2_G = .015$]. A pairwise comparison with Bonferroni correction showed that the average amplitude of electrode Pz was significantly more positive than electrode Fz [$p = .05$, $M_{Fz} = 0.39$ ($SE = 0.49$), $M_{Pz} = 1.66$ ($SE = 0.41$)].

On lateral sites the effect of CONDITION was significant [$F(3, 69) = 4.61$, $p = .005$, $\eta^2_G = .067$]. A pairwise comparison with Bonferroni correction showed that condition ***a*** differed significantly from condition ***c*** and condition ***d***. Condition ***b*** differed from condition ***c*** with marginal significance and differed significantly from condition ***d***. Finally, condition ***c*** differed significantly from condition ***d***. Table 6.2 (on page 124) summarises the means and standard errors of the four conditions and the *p*-values of the multiple comparisons that were used to determine differences between all conditions at lateral sites.

			Comparison (*p*-value)			
Condition	Mean (μV)	*SE*	***a***	***b***	***c***	***d***
a	0.41	0.32	-	1	.006	$<$ .001
b	0.32	0.30		-	.004	$<$ .001
c	1.44	0.27			-	.164
d	2.39	0.28				-

Table 6.1: Means and standard errors of the amplitudes and *p*-values of the Bonferroni-corrected multiple comparisons of the four test conditions in the 300-480 ms time window at lateral sites.

			Comparison (*p*-value)			
Condition	Mean (μV)	*SE*	***a***	***b***	***c***	***d***
a	-0.59	0.32	-	1	.027	$<$ .001
b	-0.58	0.34		-	.062	$<$ .001
c	0.47	0.21			-	.014
d	1.75	0.32				-

Table 6.2: Means and standard errors of the amplitudes and *p*-values of the Bonferroni-corrected multiple comparisons of the four test conditions in the 530-790 ms time window at lateral sites.

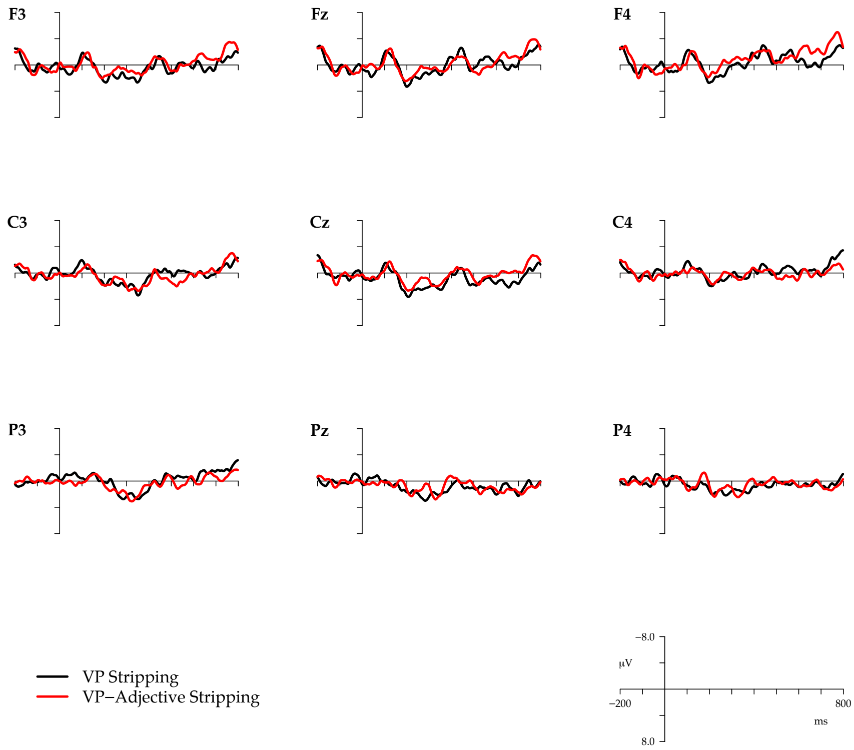


Figure 6.4: Grand averages of VP Stripping condition (***a***) and VP-Adjective Stripping condition (***b***) at onset *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 120.

6.2.5 Discussion

In the current experiment, no difference in comprehension accuracy between conditions could be established, while at the same time the factor CONDITION related to significant effects in the electrophysiological data. In that sense, this is an example of diverging results of different measurement types as discussed in Chapter 4.1.3. Based only on offline behavioural data, I would have concluded that interpretation of Stripping does not depend on the amount of structure to be retrieved and integrated, which is in line with a cost-free resolution mechanism, as is predicted by both a Copy α and cue-based accounts. However, the online data indicate that the parser does require additional ef-

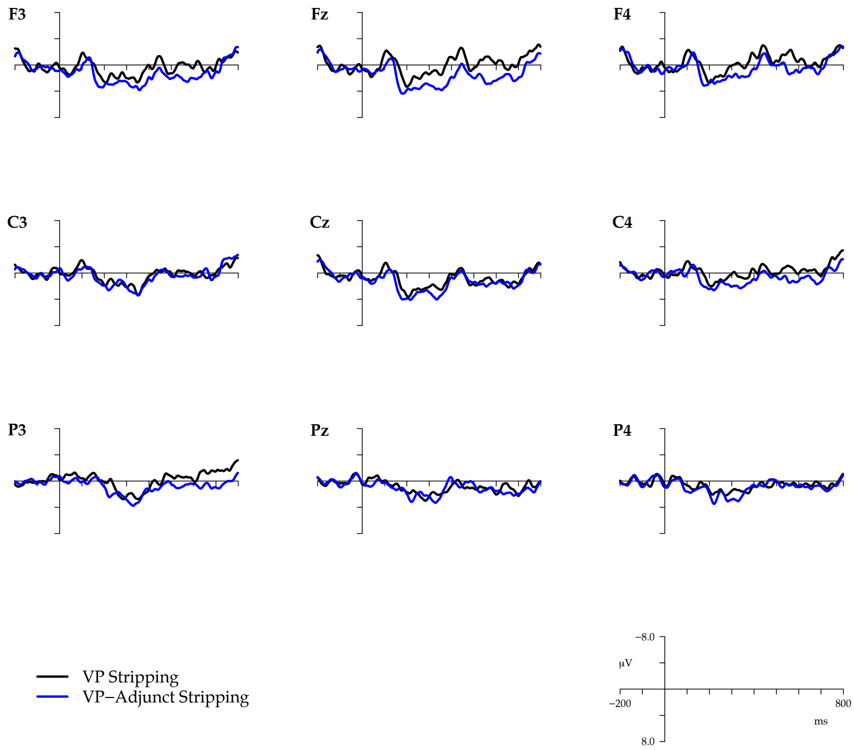


Figure 6.5: Grand averages of VP Stripping condition (***a***) and VP-Adjunct Stripping (***c***) at onset *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 120.

fort as a function of the size of the structure to be retrieved and integrated. As was suggested in the post-hoc analysis of the experiment described in section 6.1.5, the pattern of two consecutive positive deflections may be interpreted as reflecting a two-stage resolution mechanism.

In both time windows, the factor CONDITION yielded significant effects, most prominently at lateral sites. Although the interaction of CONDITION with ANTERIORITY was only slight, it appeared that in the first time window frontally distributed electrodes showed relatively more positive amplitudes in conditions in which the adjunct needs to be retrieved (conditions ***c***-***d***). Because of this frontal distribution, it is tempting to connect the positivity to a P300, in particular the so-called P3a, that is normally related to general frontal lobe attention mechanisms possibly subserving attentive language comprehension, since the manipulation in question is linguistic. Although only the behavi-

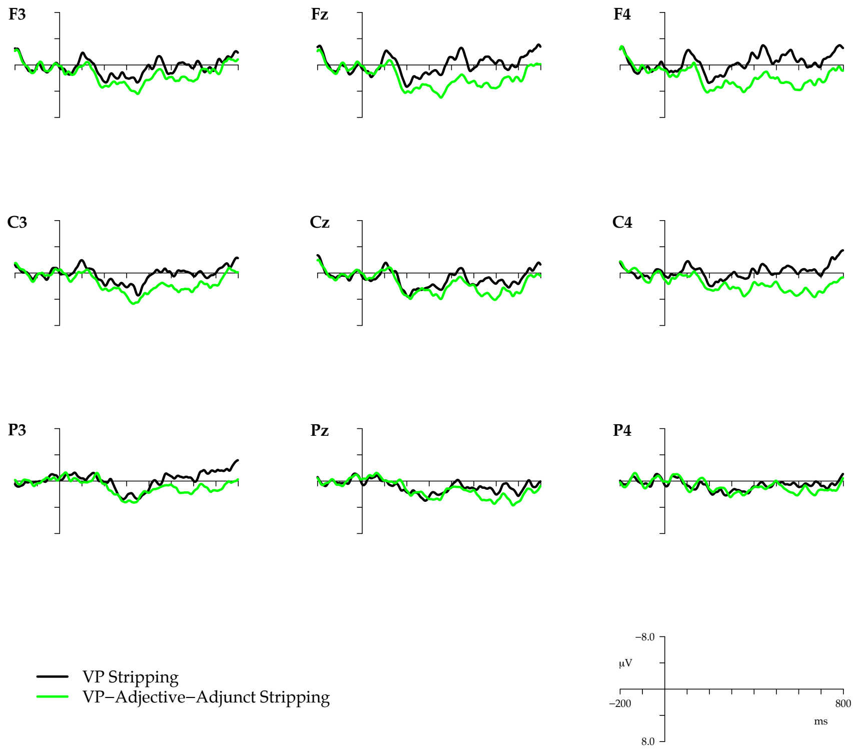


Figure 6.6: Grand averages of VP Stripping condition (***a***) and VP-Adjective-Adjunct Stripping (***d***) at onset *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 120.

oural data of the the working memory was analysed, it does not seem to be the case that general working memory plays a large role. On the one hand, the correlation between comprehension accuracy and scores on the working memory task was not significant. On the other hand, P300 effects that are related to working memory tend to correlate with activity in parietal brain areas (Polich, 2007).

Following Friederici (2002)'s model of auditory sentence comprehension, the frontal lobe can be implicated in working memory processes, but these are not reflected by positive deflections. An ELAN effect is connected with memory of syntactic structure, while semantic features are usually reflected by negative deflections between 300-500 ms. All of these negativities are assumed to originate from frontal areas. In general, the working memory data in this experiment follow the trend as seen in earlier tests: the visual condition is

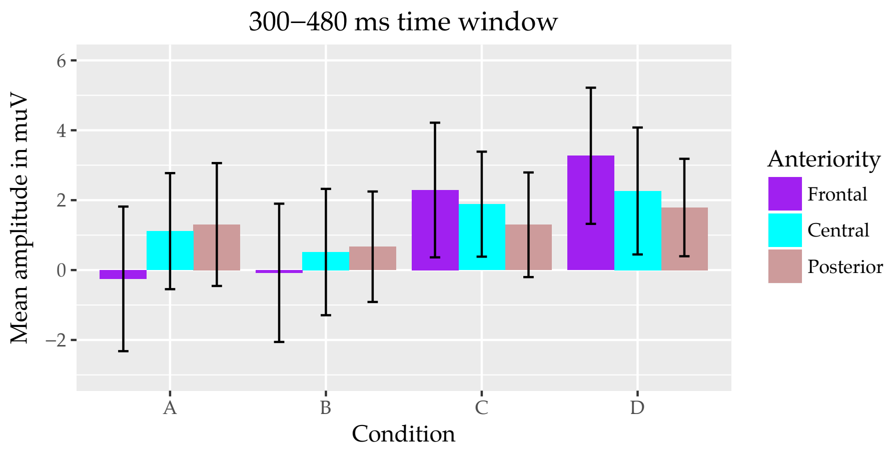


Figure 6.7: Error bar chart of the the means of voltages per condition at onset *ook* in time window 300-480 ms, grouped by anteriority on midline electrodes.

performed best, followed by random counting; the lowest scores are seen in the auditory task; the scores on the memory task barely correlate with the sentence comprehension ratios. All in all, it seems that working memory as assessed does not play a large role.

Since the critical word appeared with a comma, some part of the first positive deflection should be expected to be related to prosodic phrasing which can be measured as a CPS component (Closure Positive Shift). This seems on a par with Drury et al. (2016)'s study that reports comma effects starting at 300 ms – particularly at frontal electrodes – after onset of the critical word in sentences such as in (5).

(5) a. John said Mary was the nicest [...].
b. John, said Mary, was the nicest [...].

Typical CPS effects were shown at proper names "John," and "Mary," (presented with comma) when compared to their comma-less counterparts. However, as noted in section 6.1.6, a CPS may be sustained by acoustic as well as linguistic cues, which might explain why it can be modulated by adding structure that needs to be retrieved – as was done in the current experiment. Thus, the retrieval of the antecedent partly depends on prosodic, and possibly, syntactic and semantic information types.

It remains difficult to understand the exact nature of linguistic information types that are retrieved. That is, the relative import of these types is still unclear. Before I started the ERP experiments and in line with Friederici (2002) mentioned above, I expected to find effects related to the retrieval of a fully-fledged syntactic structure early on, possibly in the form of an ELAN component. However, again, I cannot be sure whether such an effect was apparent since it might have been cancelled out by a CPS and, possibly, other P300-related processes. Therefore, associating the early positivity directly with a Copy α mechanism would be tenuous.

If an early syntax-related component is not found, this does not necessarily entail that no structural information is retrieved. With respect to linguistic information types during retrieval, the early positive component could be a member of the P600 family. It has been argued that the latency of a P600 may vary as a function of ease of retrieval. For example, (Gouvea et al., 2010:175) claim that P600 may occur quite early if "retrieval processes that are needed to initiate structure building" – such as during the resolution of *wh*-dependencies – is relatively easy. Interestingly, such dependencies elicited a positivity between 300-500 ms that had initially a more anterior scalp distribution before shifting to a more 'standard' posterior distribution reflecting integration processes. Yet, the retrieval of an object *wh*-phrase is qualitatively different from retrieving elided structure. A *wh*-phrase is kept in memory pending integration after which it can be fully interpreted, whereas an elided phrase needs to be searched in memory to be recovered as it has already been fully interpreted. In other words, contrasting with ellipsis resolution, if a listener encounters a *wh*-phrase, the parser 'knows' that it will be needed later on.

Recall that Kaan et al. (2004) also found a fronto-central positivity between 300-500 ms during the resolution of Gapping in English, which they cautiously related to the retrieval of the preceding verb information, without being clear as to the specifics of this information. The retrieval processes converge as to their timing and although syntactic information may be necessary to proceed to the integration phase, it does not necessarily mean that this information amounts to a fully-fledged structural representation at the point of retrieval.

Given that we may regard the early positivity to be connected to a P600, it is notable that a P600 has been related to both syntactic processes and semantic processes, ranging from processes of syntactic revision and repair (for example Friederici et al., 2002), syntactic integration (for example Kaan et al., 2000) to semantically related mechanisms (see for example Burkhardt, 2007). It has been proposed by Kaan and Swaab (2003) that late frontal positivities are triggered by "discourse complexity" whereas posterior positivities relate to the repair of ungrammatical structure. While an interaction between CONDITION and ANTERIORITY was statistically – albeit marginally – demonstrated in the first time window on midline electrodes, in the second time window no such interaction was apparent. Possibly, prompted by a prosodic break, "discourse complexity" was observed by the processor already early on. That is, a more fully interpreted chunk is analysed upon retrieval.

While the early component seems to be an amalgam of several (domain-general and linguistic) processes related to prosodic phrasing, attention and retrieval, the positive component as found in the second time window was predicted to the extent that the additional retrieved information would have an impact on the integration part of the process – if no complete syntactic structure was retrieved. Notably, a gradual difference between conditions in later time windows was most significant (see Tables 6.1 and 6.2 for a comparison of numbers). In line with Kaan et al. (2013)'s result, I regard these effects as related to the P600 family involved in integration processes. Since the retrieved chunks of interpreted structure consist of different information types we may conclude that the integration with the remnant structure is relatively costly.

All in all, it seems that ellipsis resolution follows an inference-by-inference parsing mechanism, during which structure is built step-by-step and processed information is integrated. Notably, as I mentioned in Chapter 3.3, Frazier and Clifton (2001) suggest that Gapping resolution might follow such a mechanism. While this incremental procedure is widely acknowledged for sentence processing (as noted in the first chapter, Section 1.3.2), a main locus of integration processes in a relatively late time window is assumed to be reflected by late positivities. In 'normal' sentences, integration hinges on retrieval of incoming words. Difficulty of retrieval of words and semantic complexity can be demonstrated by a negative deflection around 400 ms after onset (i.e. N400). This mechanism forms the basis of a recent and computationally sustained model put forward by (Brouwer & Hoeks, 2013)'s "Retrieval-Integration" account. During ellipsis processing, the retrieval of lexical items is 'replaced' by the retrieval of missing linguistic structure.

Since a possible antecedent for ellipsis is encoded in memory moments before the elision, it is conceivable that the retrieval process of the antecedent starts slightly faster than lexical retrieval. Possibly, and on a par with lexical retrieval, during retrieval of elided structure a semantic representation is targeted at first instance, but it may be steered by additional information types. As with lexical retrieval, the amplitude of the component related to retrieval is modulated by the relative complexity of "search-and-find". During the integration phase several levels are being linked. Bearing in mind that integration processes were reflected as a broadly distributed late positivity, it may be concluded that several processes work in parallel, but also, that it remains unclear what the relative weight is of each of these processes.

By and large, Copy α seems untenable to account for Gapping and Stripping. However, it is still an open question to what extent structural information is analysed. For example, it could be that during integration (part of) this information is "shared" (Frazier & Clifton, 2005) or "recycled" (Arregui et al., 2006; Frazier, 2013). At the same time, the results are not wholly compatible with the predictions of the cue-based account as proposed by Martin and McElree (2008). However, we could assume that the retrieval stage is affected by the propositional content as generated by additional structure. While a clear and detailed description of cues is still lacking, in principle, different information types may all potentially be described in terms of cues. Since a cue-based account refers to more fully interpreted structure, I suggest that some cue directly points to propositions. As I have argued in Chapter 2.4.2, in the case that an adjunct is added, a secondary *proposition* becomes available, which possibly induces an additional retrieval step. Considering it to be a form of "cue overload", the additional propositional content may give rise to a discourse complexity that can be measured as a frontally distributed positivity. Coherent with an inference-by-inference parsing mechanism, after retrieval of a set of cues, the integration stage applies during which different levels of analysis are unified.

In a nutshell, although processes of retrieval and integration are visible as positive deflections, they remain unnoticed in terms of a behavioural measure of the kind I used in this experiment.

6.3 Conclusion

In two experiments, I investigated the role of the size of structure in Gapping and Stripping constructions. While the design of the first experiment was not optimal, I was able to interpret the collected data in a post-hoc manner, being very cautious as to the conclusions I drew. Since a comma was present at the critical measure point in the alternative analysis, a major role of prosodic phrasing was apparent as reflected by a CPS. Furthermore, the amount of new information in the right conjunct – the clause in which the amount of structure was modulated – differed between conditions. Therefore, it was not possible to make a straightforward comparison between conditions and a follow-up experiment was proposed. Provisionally, I concluded that the recovery of elided structure starts at around 300 ms after onset of the critical word and is reflected by positive deflections; furthermore that a secondary positive component can be demonstrated signalling more complex integration processes.

In the second experiment, I controlled for the amount of new information in the right conjunct and manipulated the amount of structure to be deleted in the left conjunct. Again, the analysis was influenced by a comma that was presented with the critical word and a CPS-like component was found. However, since its amplitude was modulated by the amount of structure to be retrieved I suggested that other processes, linguistic as well as domain-general (attentional, evaluative, or memory mechanisms), are at work too. Additionally, a secondary positive deflection was apparent in conditions with relatively large deletions. In line with the findings of the first experiment as well as the replicated study in Chapter 5.1, I suggested that the secondary (late) positivity relates to a P600 and reflects the relative difficulty of integration of additional propositional content.

The main goal of the experiments was to differentiate between two models of ellipsis resolution that I discussed in Chapter 3.3 and which I suggested predictions of in Chapter (1). I hypothesised that under a Copy α account, manipulation of gapped structure size would be reflected by early ERPs related to retrieval of a fully-fledged syntactic structure. Alternatively, I suggested that a cue-based account predicts relative ease of retrieval but a relatively more costly integration process. With the current data, I was able to show that both retrieval and integration processes may be affected by modulation of structure. However, it does not seem to be the case that retrieval processes are only driven by syntax-related mechanisms. I suggest that at least (subvocalic) prosodic phrasing and processes related to attention are involved. I have further argued that a discourse complexity might be noticed as early as the retrieval starts, that is, retrieval processes target more fully interpreted chunks. In the next chapter I further investigate the import of semantic information types during retrieval and integration.

As it stands, it seems that a cue-based account should be able to account for the findings in this chapter, though a clear description of how exactly a cue should be formulated is still unclear. What is clear is that the current data are supported by theoretical accounts that acknowledge the multidimensional character of ellipsis resolution, which I have discussed at length in Chapter 2.4. Future research is needed to investigate how theoretical constructs may feed into a processing account based on cues.

CHAPTER 7

ERP experiment III: Semantic complexity

This chapter reports an ERP experiment on the difference between determiner *de* “the” and quantifiers *elke/alle* “every / all” in Stripping constructions. I thank Isabella Jordanoska for assistance during data collection.

7.1 Introduction

In the experiments in the previous chapter it appeared that manipulations of structural complexity did not elicit an ERP (exclusively) related to syntactic working memory processes. It seems that retrieval mechanisms are partly dependent on prosodic and attentional processes and possibly semantic representations – the propositional content. Also, the manipulation of the linguistic structure to be retrieved appeared to impact the (proposed) integration phase. To further investigate retrieval and integration processes, an experiment was designed to focus on the semantic aspect. As I have pointed out in Chapter 2.4.2, quantifying expressions may be a burden on mechanisms of movement and/or copying since additional structural information has to be analysed. Extending this insight to processing, a Copy α account predicts a structural processing cost during recovery of ellipis in which quantification is contained. For example, a processing cost is expected in Stripping constructions when measured at "too" in (1b), when compared to (1a). If the recovery process is contingent of the retrieval of a full-fledged syntactic structure, this should be reflected as a syntax-related ERP.

(1) a. Mira peeled the oranges for breakfast, and Erica too.
b. Mira peeled all oranges for breakfast, and Erica too.

In so far as a pointer can be understood as direct reference to a proposition, a cue-based account predicts a relative ease of retrieval of a proposition that involves quantification. To the extent that such a small effect can be measured, it is expected to be reflected as part of the early positivity as found in the previous chapter. It is further expected that ERPs relating to the integration phase may be undetectable if the modulation of structure (and the inherently propositional content) is relatively small.

7.2 Methods

Test materials

From the data set described in Chapter 5.3, forty-two stimulus pairs were chosen from Stripping conditions *d* and *e*. These are repeated here in (2) as conditions ***a*** and ***b***). Condition *a* can be contrasted with condition *b* to estimate the difference between a determiner and a quantifier in Stripping constructions. As explained, items within such a pair maximally differed 1.25 in their average acceptability scores. The range of average scores among chosen items was 4.38-6.50 and the means of scores did not differ between conditions [$t(41) = 1.41, p = .166, d = .218$].

(2) a. Mira schilde de sinaasappels voor het ontbijt, en Erica **ook**.
Mira peeled the oranges for the breakfast and Erica too
'Mira peeled the oranges for breakfast,
and Erica too. ' (***Determiner Stripping***)

b. Mira schilde alle sinaasappels voor het ontbijt, en Erica **ook**.
Mira peeled all oranges for the breakfast and Erica too
'Mira peeled all oranges for breakfast,
and Erica too. ' (***Quantifier Stripping***)

From the remaining pretested item sets, 21 items from the first condition and 21 from the second condition of item set (8) in section 5.3 were chosen as fillers to prevent participants from expecting a certain type of ellipsis in the second conjunct. For a complete list of test sentences and fillers and average acceptability scores see Appendix B. Note that the measure point of interest in the test sentences is *ook*. If only sentences of the Stripping type in conditions ***a*** and ***b*** were used, the participants will start to predict at the point of the word *en* how the sentence will proceed – probably resolving the ellipsis before the moment *ook* appears. Therefore, Gapping conditions as shown in (3) were added as 'Related Fillers' (a duplet is used by means of an example).

(3) a. Koen verving de kast in de woonkamer, en Judith de lamp in de gang.
Koen replaced the cabinet in the living room and Judith the lamp in the hall
'Koen replaced the cabinet in the hall, and Judith the lamp in the hall.' (*Determiner Gapping*)

b. Koen verving elke kast in de woonkamer, en Judith de lamp in de gang.
Koen replaced every cabinet in the living room and Judith the lamp in the hall
'Koen replaced the cabinet in the hall, and Judith the lamp in the hall.' (*Quantifier Gapping*)

Filler sentences were selected in such way that the differences in rating means did not differ between the two filler sentence types (Determiner Gapping and Quantifier Gapping) as a paired t-test showed [$t(20) = 0.35, p = .73, d = .076$]. An additional set of 42 unrelated fillers were added. Between fillers and test conditions, the difference in means was kept as small as possible. A one-way ANOVA was run to establish the difference of the means between four groups (Determiner Stripping, Quantifier Stripping, Related Fillers and Unrelated Fillers) each consisting of 42 items [$F(3,\ 164) = 1.61, p = .19, \eta^2 = .028$]. Table 7.1 summarises the descriptive statistics of the selected stimuli. The two test conditions were divided over two lists and combined with the related and unrelated fillers in such a way that only one item per test pair was presented once to each participant.

Condition	Mean	N	Std. Error
Determiner Stripping	5.46	42	0.10
Quantifier Stripping	5.32	42	0.08
Related filler	5.52	42	0.08
Unrelated filler	5.55	42	0.06
Total	5.46	168	0.04

Table 7.1: Means of rating of selected test sentences and fillers.

Participants

Twenty-two right-handed native Dutch participants with normal or corrected-to-normal vision took part in this study and were paid €15. The EEG data of one participant was not recorded due to technical failure. Three others were discarded from the analysis due to too many artefacts resulting in fewer than eight trials in one of the conditions. Of the remaining 18 participants six were male and the mean age was 22.28 (range 18-28). The experiment followed the Ethics Committee regulations of the Humanities Faculty of Leiden University, which approved its implementation. Participants gave informed consent before the study.

Procedure

Participants were comfortably seated in a dimly lit sound-proof room at a distance of approximately 90 cm of a 19 inch LCD monitor. One-hundred-and-eight test sentences were presented in a random order using the presentation software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The sentences were presented word by word in Verdana font (36pt). Each word was presented for 300 ms with a 300 ms fixation cross interval. Presentation of a trial started with a fixation cross for 1,000 ms. Every sentence was followed by a content question to encourage comprehension. A blank screen with a duration of 1,500 ms intervened between the last word of a trial and the presentation of the comprehension question. For half the participants the left response button referred to "YES", for the other half the left button referred to "NO". Participants were given a break after 12 sentences and could proceed at their own pace. The comprehension questions referred to different parts of the sentences equally. Before the actual test, the participants were able to get used to the task with four practice sentences.

The experiment was concluded with a working memory test as used before.

In total, the experiment took about 1.5 hours per participant, including set-up.

Apparatus and electrophysiological recording

A description of the recording set-up can be found in Chapter 5.1.1.

Data analysis

Using Brain Vision Analyzer Version 2.0 (Brain Products, Munich, Germany) the EEG data were preprocessed before analysis to reduce noise and artefacts as much as possible. Eye blinks were corrected using an Independent Components Analysis procedure (Makeig, Bell, Jung, & Sejnowski, 1996). Remaining artefacts were rejected on the basis of the same criteria as used in all earlier experiments reported in this thesis. Epochs of 1,000 ms were computed with a 200 ms pre-stimulus baseline and ERP grand averages were time-locked to the onset of the target words *de* and *elke*. 14.15% of the trials were excluded from the analysis; of the 21 trials presented per condition, an average of 18.03 ($SE = 0.73$) were retained.

In the current experiment, a repeated measures ANOVA was planned using within-subjects factors CONDITION (two levels: Determiner Stripping, Quantifier Stripping), ANTERIORITY (3 levels: frontal, central, posterior), and, for analyses involving lateral sites, HEMISPHERE (2 levels: left, right).

7.3 Behavioural results

On average the accuracy on the comprehension questions of the two test conditions was 96.70% ($SE = 0.65\%$). The accuracy scores were similar for both test conditions [$M_{\text{Determiner}} = 96.83\%$, $M_{\text{Quantifier}} = 96.56\%$] as the difference in mean values was not significant as shown by a paired t-test on the scores [$t(17) = 0.19$, $p = .848$, $d = .046$]. The accuracy scores of the test conditions (Determiner Stripping and Quantifier Stripping) were further compared to the related filler conditions (Determiner Gapping and Quantifier Gapping). A significant difference was apparent as shown by a repeated measures by subjects on the scores [$F(3, 51) = 17.77$, $p < .001$, $\eta^2_G = .366$]. As can be seen in Table 7.2, which reports the results of a multiple comparisons procedure with Bonferroni correction, mean comprehension accuracy differs significantly between Stripping and Gapping conditions but not within these conditions.

			Comparison (p-value)			
Condition	Mean ratio	*SE*	**D Strip**	**Q Strip**	**D Gap**	**Q Gap**
D Strip	.968	.010	-	1	$<$ .001	$<$ .001
Q Strip	.966	.010		-	$<$ .001	.005
D Gap	.876	.016			-	1
Q Gap	.881	.019				-

Table 7.2: Means and standard errors of the accuracy scores and p-values of the Bonferroni-corrected multiple comparisons of the test conditions (**D**eterminer **Strip**ping and **Q**uantifier **Strip**ping) and related filler conditions (**D**eterminer **Gap**ping and **Q**uantifier **Gap**ping).

The accuracy of the three test sessions of the working memory task was on average 70.00% ($SE = 2.79\%$). Per condition, the scores were: $M_{\text{Random Counting}} = 67.78\%$, $M_{\text{Auditory Presentation}} = 58.89\%$, $M_{\text{Visual Presentation}} = 83.33\%$. A repeated measures ANOVA by subjects with CONDITION as independent factor and ACCURACY OF NUMBER RECALL as dependent variable showed that the scores differed between conditions [$F(2, 34) = 7.75, p < .043, \eta^2_G = .202$]. A multiple comparison with Bonferroni correction showed that the visual condition differed marginally from the random condition [$p = .055$] and significantly from the auditory condition [$p < .001$].

The scores from the sentence comprehension task were compared with the scores from the working memory task. A large and significant correlation was found between the variables ACCURACY OF SENTENCE COMPREHENSION and ACCURACY OF NUMBER RECALL [$r = .632, p = .005$].

7.4 Electrophysiological results

Figure 7.1 depicts the grand averages of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at the critical measure word *ook*. Most prominently, and mainly at frontal and central electrodes, a positive deflection starting around 300 ms after onset is apparent in both conditions. At some electrodes, condition ***b*** seems to deviate from condition ***a*** in a more positive direction. However, a permutation test per sample at every electrode with independent factor CONDITION did not yield significant time windows to be analysed any further.

Since the correlation between accuracy of sentence comprehension and the scores on the working memory task was significant in this experiment, its relevance was explored. First, the relation between working memory scores and ERPs were taken into account, and then, the relation between sentence comprehension and ERPs.

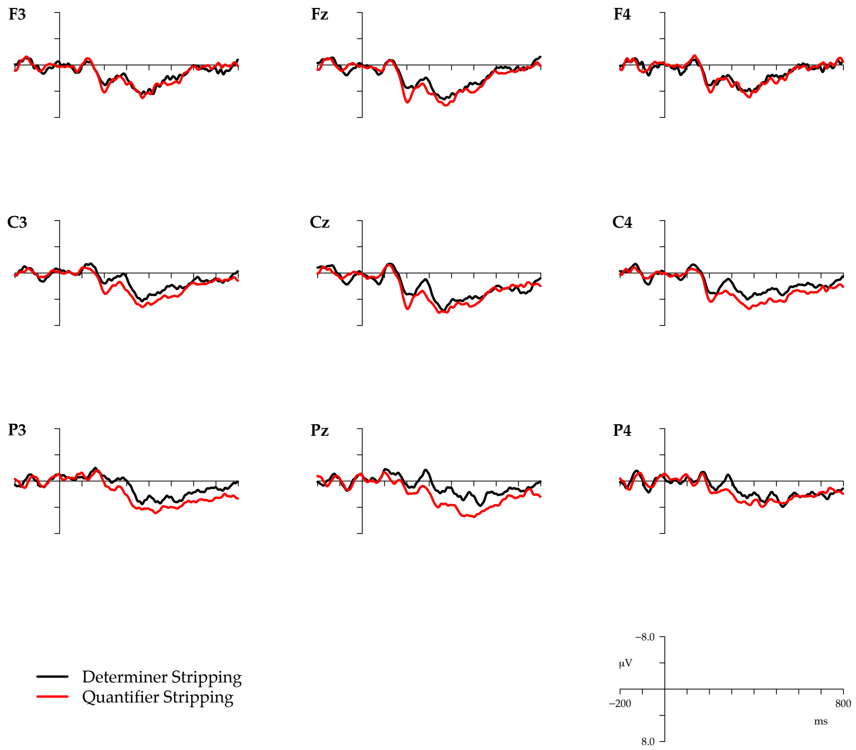


Figure 7.1: Grand averages of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

7.4.1 Relation between working memory and ERPs

On the basis of a median split, the participants were divided into two groups, (*i*) consisting of participants with relatively low accuracy on the working memory task (range: 40.00%-66.67%, $M = 57.08\%$, $SE = 2.96\%$) and (*ii*) consisting of participants with relatively high scores (range: 73.33%-93.33%, $M = 82.96$, $SE = 1.96\%$). Figure 7.2 depicts the grand averages of the first group while the second group is shown in 7.3. In each group, a permutation test per sample at every electrode with independent factor CONDITION was conducted. No effects were found. I also wished to explore the possibility of overall processing differences between groups, by taking the average across conditions and comparing it between the two groups. The result can be seen in Figure 7.4. A permutation test per sample at every electrode with independent factor GROUP did not show any effect.

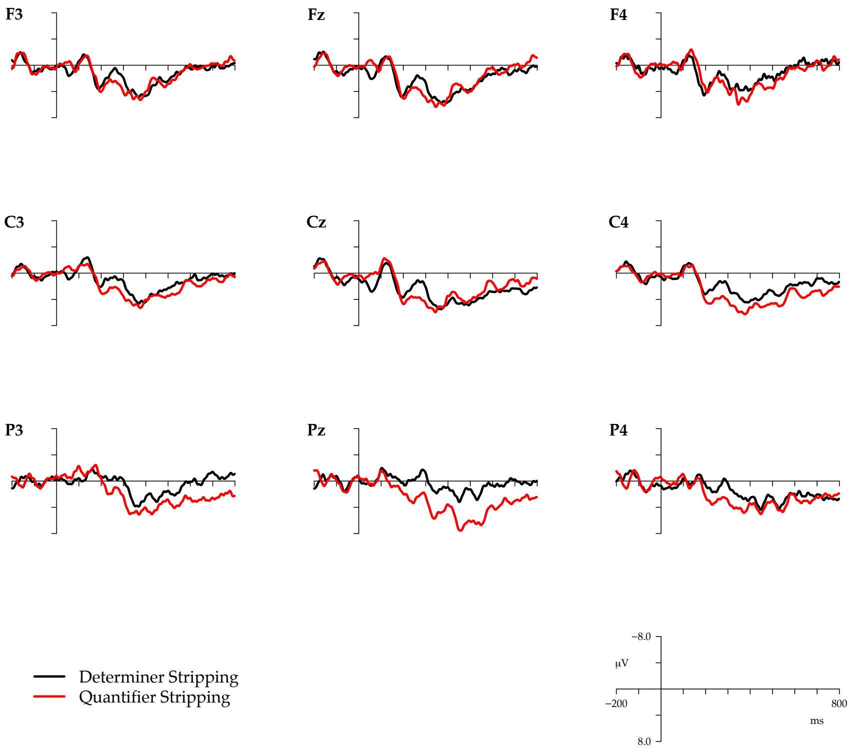


Figure 7.2: Grand averages of participants with low working memory scores (n = 9) of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

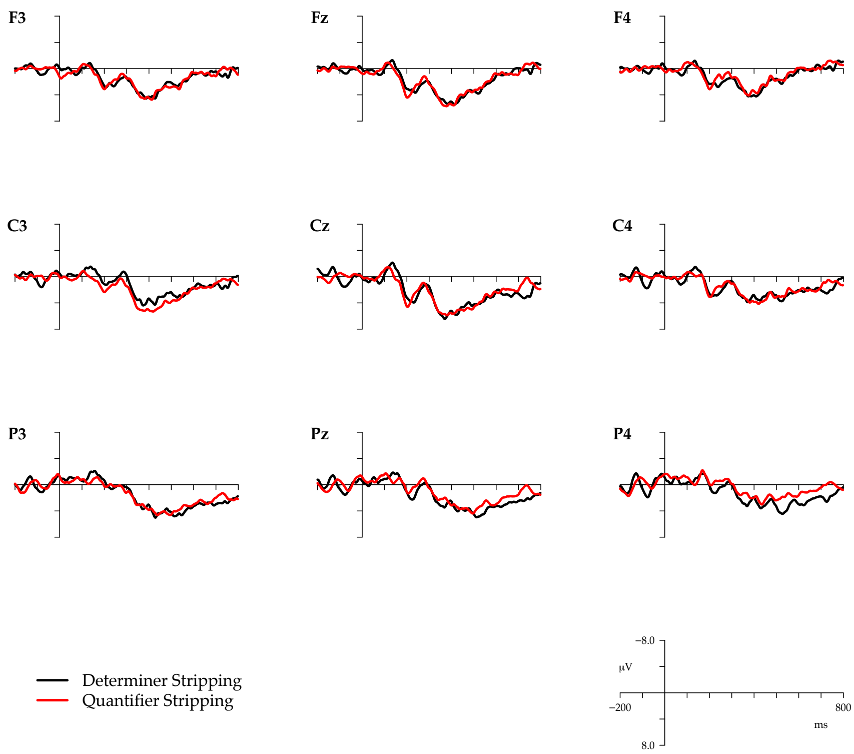


Figure 7.3: Grand averages of participants with high working memory scores (n = 9) of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

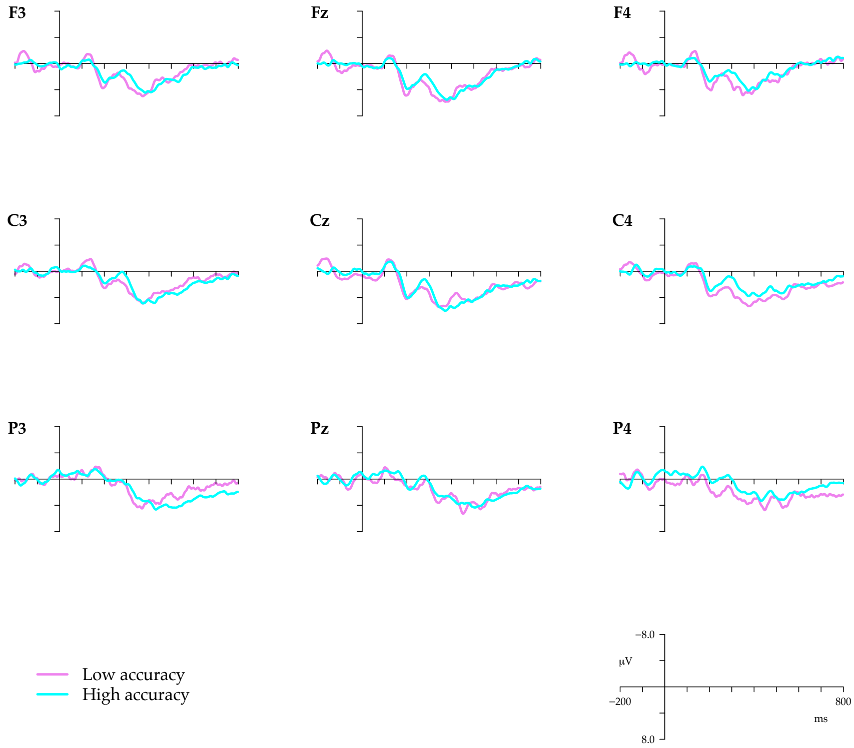


Figure 7.4: Collapsed grand averages of conditions ***a*** and ***b*** of two groups split by working memory scores: 'Low accuracy' (*n* = 9) and 'High accuracy' (*n* = 9) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

7.4.2 Relation between sentence comprehension and ERPs

Again, on the basis of a median split, the participants were divided into two groups, (*i*) consisting of participants with relatively low accuracy on the sentence comprehension task (range: 82.54%-88.89%, $M = 87.48\%$, $SE = 0.67\%$) and (*ii*) consisting of participants with relatively high scores (range: 89.68%-96.83%, $M = 93.65$, $SE = .78\%$). Figure 7.5 depicts the grand averages of the first group while the second group is shown in 7.6. In each group, a permutation test per sample at every electrode with independent factor CONDITION was conducted. No effects were found. Also, the possibility of overall processing differences between groups was explored, by taking the average across conditions and comparing it between the two groups. The result can be seen in Figure 7.7.

Based on a secondary permutation test per sample at every electrode with independent factor GROUP, two repeated measures ANOVAs were run for a time window between 340-800 ms with within-subjects factors CONDITION, ANTERIORITY and HEMISPHERE (at lateral sites), and between-subjects factor GROUP. The means differed between groups significantly on midline sites [$F(1, 16) = 5.28$, $p = .004$, $\eta^2_G = .086$] and on lateral sites [$F(1, 16) = 7.75$, $p = .001$, $\eta^2_G = .164$]. Further, an effect of ANTERIORITY was apparent on lateral sites [$F(2, 32) = 11.31$, $p < .001$, $\eta^2_G = .127$]. A post hoc multiple comparison with Bonferroni correction showed that the means of amplitudes at central electrodes were equally significantly more positive than frontal and posterior electrodes [$p < .001$, $M_{\text{Central}} = 2.99$ ($SE = 0.33$), $M_{\text{Frontal}} = 1.29$ ($SE = 0.21$), $M_{\text{Posterior}} = 1.51$ ($SE = 0.25$)].

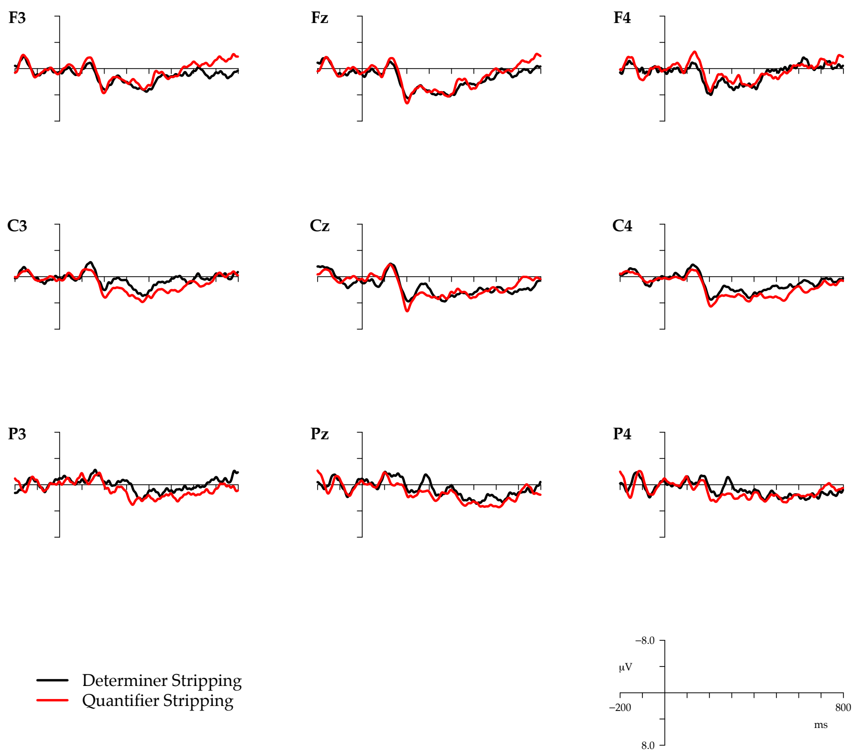


Figure 7.5: Grand averages of participants with low sentence comprehension scores (n = 9) of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

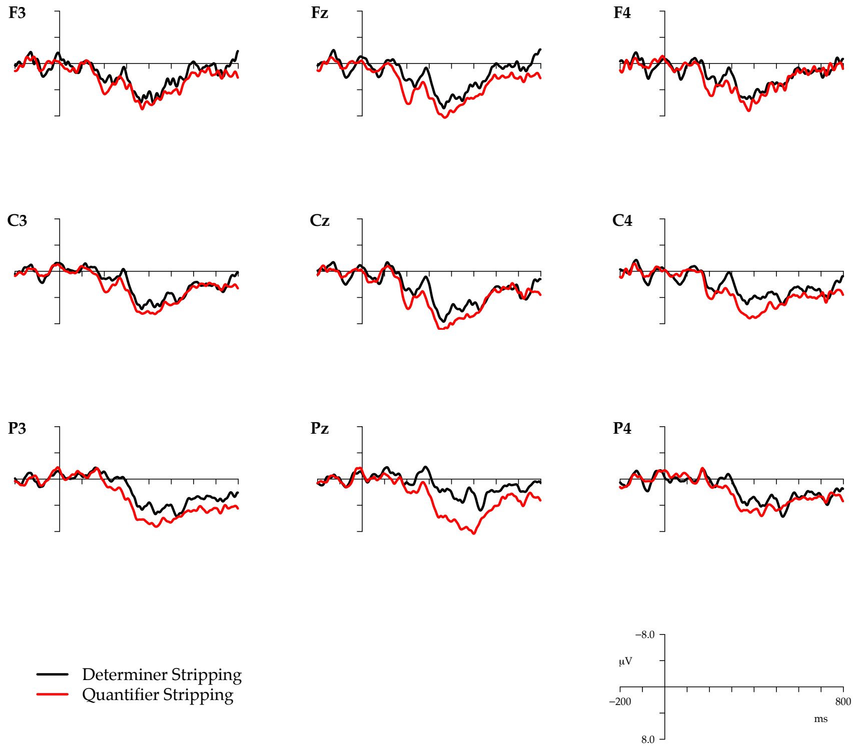


Figure 7.6: Grand averages of participants with high sentence comprehension scores (n = 9) of Determiner Stripping condition (***a***) and Quantifier Stripping condition (***b***) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

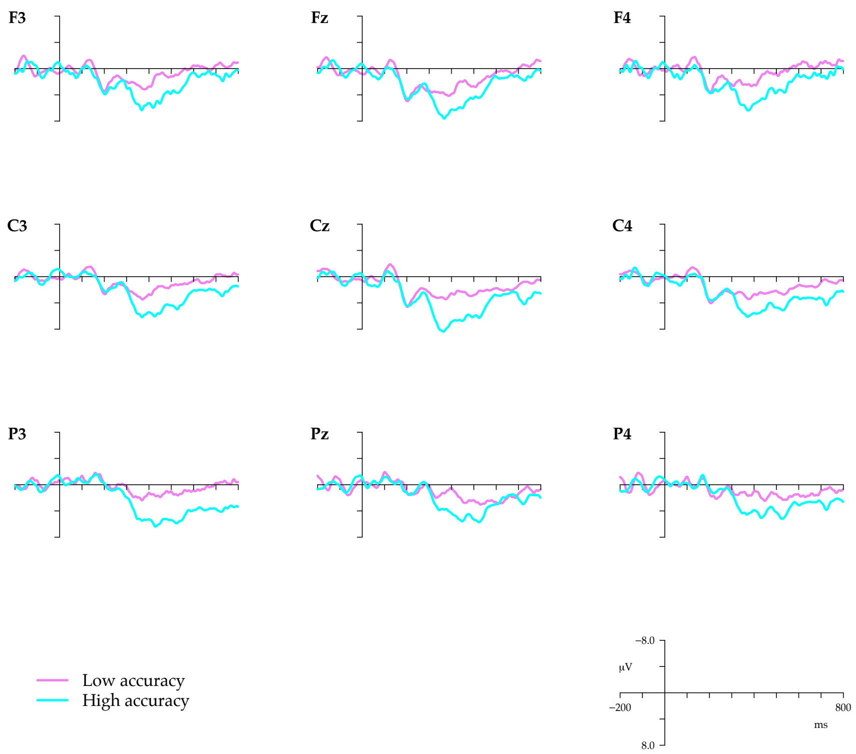


Figure 7.7: Collapsed grand averages of conditions ***a*** and ***b*** of two groups split by sentence comprehension scores: 'Low accuracy' (n = 9) and 'High accuracy' (n = 9) at onset (y-axis) *ook* at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Corresponding example sentences can be found on page 135.

7.5 Discussion

In this experiment, I investigated the difference between determiner *de* "the" and quantifiers *elke/alle* "every/all" in Stripping constructions. I hypothesised that if the recovery process is contingent on the retrieval of a full-fledged syntactic structure, this should be reflected as a syntax-related ERP. Contrastingly, I hypothesised that a pointer can be understood as direct reference to a proposition and that, as a consequence, a cue-based account should be able to account for a relative ease of retrieval of a proposition that involves quantification. Further, I expected that this retrieval process to be reflected as part of the early positivity as found in the previous chapter. I also expected that ERPs relating to the integration phase might be undetectable if the manipulation of structure (and the inherently propositional content) is relatively small.

The comprehension scores of the sentences in the main test conditions (Stripping) were at ceiling. Although rating means of these sentences did not differ from those of the related filler sentences (Gapping), the Stripping sentences were easier to comprehend. Note, that the rating means as collected in the pretest were based on the structure as well as the interpretability. While in the first experiment on structural modulations described in Chapter 6.1 no difference in comprehension scores was apparent between the Gapping conditions (***b***-***c***) and the Stripping condition (***d***), in the current experiment, it appeared that Gapping sentences were relatively more difficult to understand. Since in the current experiment rating means of the presented items were better matched than in the first experiment on structural modulations reported in Chapter 6.1, it seems reasonable to conclude that the current comprehension difference follows from the fact that in the Gapping conditions there are three contrasting phrases instead of one, thus, adding more information load to the utterance.

There was a tendency for positive deflections, starting around 300 ms after onset of the critical word in the grand averages of the whole group of participants in both conditions. At some electrodes, the positive deflection of the condition with the elided quantifier seemed larger as compared to the condition with the determiner. However, the positivity did not yield a significant difference. Meanwhile, a significant correlation was found between sentence comprehension scores and the working memory task scores (in contrast to the previous three ERP experiments reported in this thesis). Participants with high scores on the working memory task generally had high scores on the sentence comprehension task. I explored to what extent the working memory scores and sentence comprehension accuracy might be related to the ERP results.

While no effects could be established with respect to the relation between working memory scores and the ERPs found in the sentence reading task, it appeared that the positivity in both conditions as observed in Figure 7.1 was mainly generated by the group of participants with high scores on the sentence comprehension task (c.f. Figure 7.7). Considering the fact that in both test conditions a recovery process of Stripping was involved, the conditions were collapsed and compared between groups. It appeared that the group with high accuracy on sentence comprehension showed a sustained positivity starting at 340 ms. Possibly, with a larger group of participants it may be demonstrated that this sustained positivity actually consists of two positive components as found in earlier experiments. If these results can be replicated with a larger group, the question is why there is a difference in processing strategies. As discussed earlier, the positivity may be related to several mechanisms, including processes related to attention. For the time being, I tentatively conclude that participants with relatively high comprehension accuracy were actively involved during the reading task while others were passive interpreters which may have caused slightly more difficulty during the comprehension task.

Aside from this interesting exploration, the rationale behind this experiment was to investigate to what extent quantifying expressions may help to decide between two mechanisms of ellipsis resolution. I hypothesised that quantifying expressions are a burden on a mechanism such as Copy α which would predict a structural processing cost during the recovery of the ellipsis. This should be reflected as a syntax-related ERP. However, the results seem to uphold the contrasting view that a possible antecedent for ellipsis is navigated in memory using cues which point to a more fully interpreted linguistic structure which may consist of additional information types. As a consequence, ERPs related to the integration of this structure can be measured. However, it appeared that the difference between determiners and quantifiers in stripped expressions is too small to the extent that neither the retrieval nor the integration phase are reflected by a distinct ERP. Still, additional experiments are needed to confirm this.

7.6 Conclusion

In this experiment, I used a semantic manipulation to study the processing of Stripping constructions. To do so, I compared the determiner *de* "the" with quantifiers *elke/alle* "every/all" in Stripping constructions, comparing ERPs at the critical measure point *ook*.

On the hypothesis that Copy α predicts a structural processing cost during the recovery of the quantified structures, a syntax-related (early) ERP was expected as a reflection of this mechanism. Although visually a slight difference was apparent in ERPs, the difference between the determiner and quantifier conditions was not significant. Additionally, the timing and polarity of the deviance was comparable with the early positive component as found in previous experiments, on the basis of which I concluded that retrieval processes are not exclusively steered by syntax-related mechanisms. It seems that this conclusion also applies to the current experiment.

Related filler sentences with Gapping constructions were presented and all sentences were matched in terms of acceptability. As a consequence, a straightforward comparison could be made as to the comprehension scores of Stripping versus related Gapping conditions. This comparison indicated that Gapping sentences are relatively more difficult to understand than Stripping constructions.

CHAPTER 8

ERP experiment IV: Prosodic cues

In this auditory EEG experiment, the character of prosody in relation to the prediction of gapped structure was investigated. I thank Judith Kelholt for her assistance with organising and preprocessing of the stimuli, and recording of the EEG data. I am further grateful to Johanneke Caspers for recording the stimulus material. Cesko Voeten, Jos Pacilly and Olga Kepinska helped me with scripting preprocessing steps for the audio files and EEG analysis.

8.1 Introduction

So far, I have presented data from word-by-word reading tasks. In this experiment, I will investigate Gapping in the auditory modality. As discussed in earlier chapters, it has been suggested that Gapping is licensed by a discourse constraint which is dictated by syntax. Furthermore, prosodic parallelism has been put forward as an important grammatical constraint. I have argued, however, that it may be possible that prosody interacts with interpretation regardless of syntax. In the ERP experiment reported in this chapter, I tested to what extent listeners are guided by prosody during the interpretation of Gapping by comparing conditions that only differ in the way prosody is expressed.

In Chapter 2.5, I asked to what extent a sentence such as (1b), as compared to (1a), is ungrammatical, since the object in the right conjunct lacks an accented counterpart in the left conjunct in (1b). The unaccented object in the first conjunct may be a sign that it will be a candidate for deletion as is the case in (1c).

(1) a. De MAN kocht een BOEK in LONDEN, en de VROUW een KRANT in LEIDEN.
'The man bought a book in London, and the woman a newspaper in Leiden.'
b. De MAN kocht een boek in LONDEN, en de VROUW een KRANT in LEIDEN.
'The man bought a book in London, and the woman a newspaper in Leiden'
c. De MAN kocht een boek in LONDEN, en de VROUW in LEIDEN.
'The man bought a book in London, and the woman a newspaper in Leiden'

It may therefore be hypothesised that the prosody of the first conjunct predicts the remaining structure in the second conjunct. The grammar requires that the object in the second conjunct in (1b) is deleted, if the object in the first conjunct, *boek*, bears no contrastive accent (c.f. the felicity condition on contrastive fragments Griffiths & Lipták, 2014). If, however, this object is accented, as is the case in (1a), the grammar requires a parallel phrase to be apparent in the second conjunct contrasting with its counterpart. In processing terms, this state of affairs could be translated into a cue to predict upcoming (deleted) structure. In other words, a processing difficulty is predicted if an object is deaccented in the first conjunct but accented in the second conjunct. Behaviourally, a processing difficulty may be measured in terms of comprehension scores which may become lower as a function of processing difficulty.

The onset of the determiner of the object determiner phrase in the right conjunct will be taken as measure point to compare ERPs. This is comparable to the replicated ERP experiment reported in Chapter 5.1. If people predict deleted structure, but encounter a remnant instead, I expect a LAN-like component between 200-400 ms may be found on the determiner. In line with Kaan et al. (2013) and Lau et al. (2006), I will regard a LAN as an index of prediction strength (though "index of surprisal" may be more straightforward). In addition, an unexpected noun may elicit an N400 effect, although scepticism is in order given the null effect in the replication study. Additionally, auditory processing is relatively fast, with effects of unexpected items reported to start around 50 ms after stimulus onset (see for example Clementz, Barber, & Dzau, 2002).

8.2 Methods

Test materials

From the replicated study in Chapter 5.1, forty-four stimulus pairs were chosen as exemplified in (2). As explained in Chapter 5.2.2, only sentences with verbal modifiers could be included given that objects should be able to be deleted. The listener may anticipate Gapping of the verb as well as the object in (2b). In other words, this deviant condition can be compared to the control condition (2a) in which an accented contrasting element is available.

(2) a. ANOUK zond de KAART aan haar VADER, en JULIA
Anouk sent the card to her father and Julia
de BLOEMEN aan haar MOEDER.
the flowers to her mother
'Anouk sent the card to her father,
and Julia the flowers to her mother.' (***Parallel prosody***)

b. ANOUK zond de kaart aan haar VADER, en JULIA
Anouk sent the card to her father and Julia
de BLOEMEN aan haar MOEDER.
the flowers to her mother
'Anouk sent the card to her father,
and Julia the flowers to her mother.' (***Non-parallel prosody***)

Since the contrastingly accented object in the second conjunct is unexpected in (2b) as compared to the same object in (2a), an ERP effect may be expected at the determiner preceding the noun. At the determiner a negativity, possibly a LAN, may be expected (as found in the replication study), relating to the violation of a grammatical constraint. The determiner will be unexpected and therefore a LAN may be considered as an index of prediction strength. At the following noun, an N400 may be predicted since the noun is less expected in the context.

In addition to the test stimuli, ten filler sentences containing a coordination with the connective *en* "and" and twenty-two containing a coordination with the connective *maar* "but" were selected from the original data set. The experimental pairs were divided over two lists. An additional five sentences were included as practice items.

All sentences were digitally recorded (44.1 kHz) by a professionally trained Dutch native phonetician in a sound-proof room using a directional Sennheiser MKH-416 condensor microphone. Items were edited and analysed in Praat (Boersma & Weenink, 2017). The duration of experimental stimuli was 4,624 ms on average (SE = 36.08) and durations did not differ between conditions [$t(43)$ = 0.73, p = .471; M_a = 4,650, SE_a = 53.40; M_b = 4,598, SE_b = 48.88]. Representative examples of prosodic contours and their ToDI (Dutch version of ToBI, tones and breaks indices, as developed by Gussenhoven, 2005) transcription are displayed in Figure 8.1.

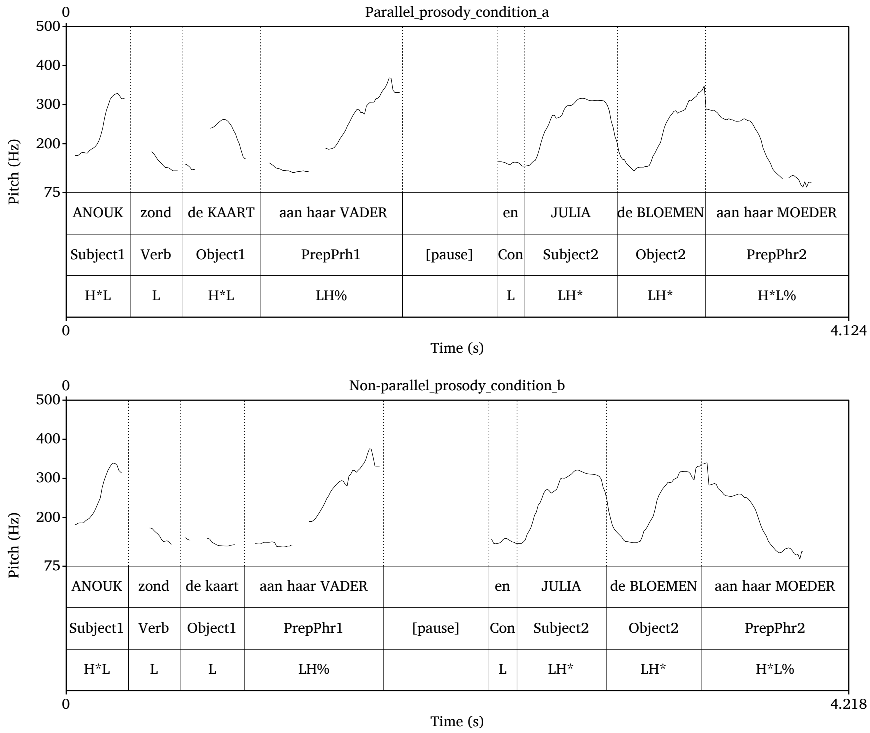


Figure 8.1: Prosodic contours of the Parallel prosody condition (***a***) are shown in the top panel and for the Non-parallel prosody condition (***b***) below. For every phrase, the content is provided, as well as a label and transcription in ToDI.

Each sentence can be decomposed into individual phrases. For every phrase (as depicted in Figure 8.1), the mean of the fundamental frequency and duration were calculated individually. Figure 8.2 represents the relative differences of the fundamental frequency between phrases and Figure 8.3 represents the differences in duration. Additional pairwise t-tests using a Bonferroni correction were performed both to establish differences between and within phrases in terms of pitch and to determine differences in means of duration. The relevant comparisons are listed in Table 8.1.

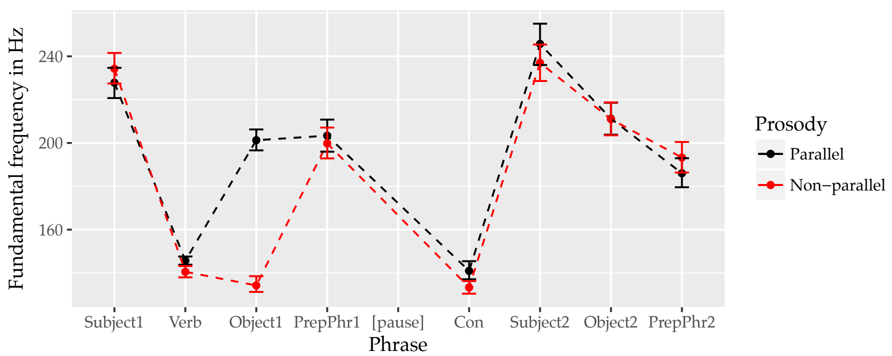


Figure 8.2: Means of fundamental frequency per phrase grouped by conditions of experimental sentences (bars indicate 95% confidence intervals).

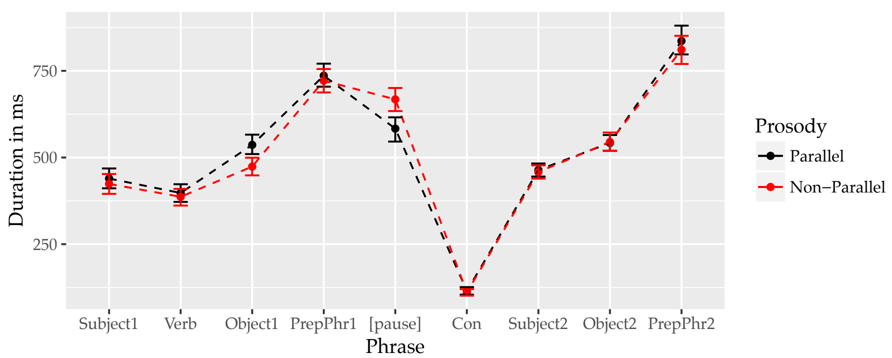


Figure 8.3: Means of duration per phrase grouped by conditions of experimental sentences (bars indicate 95% confidence intervals).

Figures 8.2 and Table 8.1 clearly show the expected differences in means of pitch between the objects of the first and second conjunct. The duration difference between the prepositional phrase of the first conjunct and its counterpart in the second is due to final lengthening. In Dutch, final lengthening of an utterance is longer than final lengthening at a phrase boundary within an utterance (see for example Cambier-Langeveld, Nespor, & Heuven, 1997). A few additional results are notable. In the Parallel prosody condition,

Phrase	Subject				Verb	
Comparison	1a-1b	1a-2a	1b-2b	2a-2b	a-b	
F0	1	.016	1	1	1	
Duration	1	1	1	1	1	
Phrase	Object					
Comparison	1a-1b	1a-2a	1b-2b	2a-2b		
F0	<.001	<.001	<.001	1		
Duration	.425	1	.108	1		
Phrase	PrepPhr				[pause]	en
Comparison	1a-1b	1a-2a	1b-2b	2a-2b		
F0	1	.129	1	1	NA	1
Duration	1	<.001	.003	1	.009	1

Table 8.1: *P*-values of Bonferroni corrected pairwise t-tests. “1” and “2” refer to the first conjunct and second conjunct, respectively; “a” and “b” refer to condition ***a*** (Parallel prosody) and ***b*** (Non-parallel prosody) respectively.

the mean of pitch of the subject in the left conjunct was significantly lower than of that of the right conjunct. In the Non-parallel condition, there was no difference in pitch. Despite this, no significant differences between conditions within phrase position were found. Lastly, it appeared that the relatively shorter durations in the first conjunct in the Non-parallel prosody condition were compensated during the pause which was longer in condition ***b***. As a result, the onset of the determiner in the second conjunct relative to the onset of the sentence was almost equal between conditions [$M_a = 3272$, $SE_a = 0.03$; [$M_b = 3243$, $SE_b = 0.03$].

Participants

Twenty native Dutch participants with normal hearing (three left-handed, three male, $M_{Age} = 23.05$, range 19-41) took part in this study and were paid €15. The responses of all participants are taken into consideration in the behavioural data analysis. The EEG data of six participants were disregarded after preprocessing because of too many artefacts. Of the remaining participants (three male, mean age = 23.93, range 19-41) two were left-handed. The experiment followed the Ethics Committee regulations of the Humanities Faculty of Leiden University, which approved its implementation. Participants gave informed consent before the study.

Procedure

The stimuli were presented in a fully randomised order using the presentation software E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Stimuli were counterbalanced in such a way that only one item per sentence pair was presented to each participant. Participants were comfortably seated in a dimly lit sound-proof room at a distance of approximately 90 cm from a computer monitor. Each stimulus was preceded by a fixation cross ("+") which appeared at the centre of the screen and remained there for 1,000 ms. Then the audio file was played. After a pause of 1,500 ms a yes/no comprehension question appeared on the screen. The questions referred to different parts of the sentences equally and participants were instructed to answer as quickly and accurately as possible. For half the participants the left response button referred to "YES", for the other half the left button referred to "NO". Between pressing the response button and the next trial, a pause intervened lasting 1,000 ms. After seven trials, a break occurred which could be ended by the participant when they were ready to proceed. Before the actual experiment which contained 76 trials in total, five trials were presented as a practice session.

The experiment was concluded with the working memory test as used in earlier experiments.

The experiment took about 1.5 hours per participant in total, including set-up.

Apparatus and electrophysiological recording

A description of the recording set-up can be found in Chapter 5.1.1.

Data analysis

Using Brain Vision Analyzer Version 2.0 (Brain Products, Munich, Germany), the EEG data were preprocessed before analysis to reduce noise and artefacts as much as possible. Eye blinks were corrected using an Independent Components Analysis procedure (Makeig et al., 1996). Remaining artefacts were rejected on the basis of the same criteria in all earlier experiments reported in this thesis. Epochs of 1,000 ms were computed with a 200 ms pre-stimulus baseline. ERP grand averages were time-locked to the onset of the target word *de*, the determiner of the object in the second conjunct. In total 19.97% of the trials were rejected. Out of 22 trials per condition per participant, 16.93 trials on average in condition ***a*** and 18.29 in condition ***b*** were retained for analysis.

To determine time-windows of interest, a permutation test was carried out at every sample for each electrode with independent factor CONDITION. Considering the small sample, no further statistical tests were run and only an exploratory analysis is carried out.

8.3 Behavioural results

Due to a scripting error one sentence pair was assigned an unrelated question. Therefore, this pair has been excluded from the behavioural analysis. No participants were rejected on the basis of accuracy (M = 96.51%, SE = 0.65%).

The accuracy scores were similar for both conditions ($M_{\text{Parallel prosody}}$ = 96.27%, $M_{\text{Non-parallel prosody}}$ = 96.76%. The difference in mean values was not significant as shown by a paired t-test [$t(19) = -0.39$, $p = .699$, $d = .088$].

The data from the working memory test of one participant were compromised due to a technical malfunction. Therefore, the following analysis is based on 19 participants. The average accuracy score of the three test sessions was 62.46% (SE = 2.87%). Per condition the scores were: $M_{\text{Random Counting}}$ = 56.84%, $M_{\text{Auditory Presentation}}$ = 60.00%, $M_{\text{Visual Presentation}}$ = 70.53%. Although numerically the difference between the random counting and visual conditions seemed large, a repeated measures ANOVA by subjects with CONDITION as independent factor and ACCURACY OF NUMBER RECALL as dependent variable yielded only marginal significance [$F(2, 34) = 2.76$, $p = .077$, $\eta^2_G = .066$].

The scores of the comprehension task of the ERP experiment were compared with the scores of the working memory task ($n = 19$). A small, non-significant correlation was found between the variables ACCURACY OF SENTENCE COMPREHENSION and ACCURACY OF NUMBER RECALL [$r = .300$, $p = .212$].

8.4 Electrophysiological results

In Figure 8.4, grand averages of Parallel prosody condition ***a*** compared to Non-parallel prosody condition ***b*** at onset of the determiner of the second object are visualised. Average onset of the noun is 141 ms (SE = 0.004) after onset of the determiner. Since an improper baseline can be observed, especially at frontal and central electrodes, further interpretation is not warranted.

In order to identify any between-condition differences that occur before the onset of the determiner, epochs of 1,375 ms were computed with a 575 ms pre-stimulus time-window (relative to the onset of the determiner) and an 800 ms time-window after target onset. The pre-stimulus time-window of 575 ms was determined on the basis of average durations of the connective *en* [$M = 113.00$, $SE = 3.87$] and the subject [$M = 460.98$, $SE = 7.16$] preceding the target. The approximate onset of the connective was taken as the start of a baseline of 100 ms. During determination of longer epochs, 24.84% of the trials were rejected. Out of 22 trials, 6.14 trials on average in condition ***a*** and 4.79 in condition ***b*** were excluded from calculating the grand averages. Figure 8.5 shows the grand averages of Parallel prosody condition (***a***) compared to Non-parallel prosody condition (***b***) of the longer epochs as well as the results of the permutation test ($p \leq .05$) that was carried out for each electrode (by the factor CONDITION).

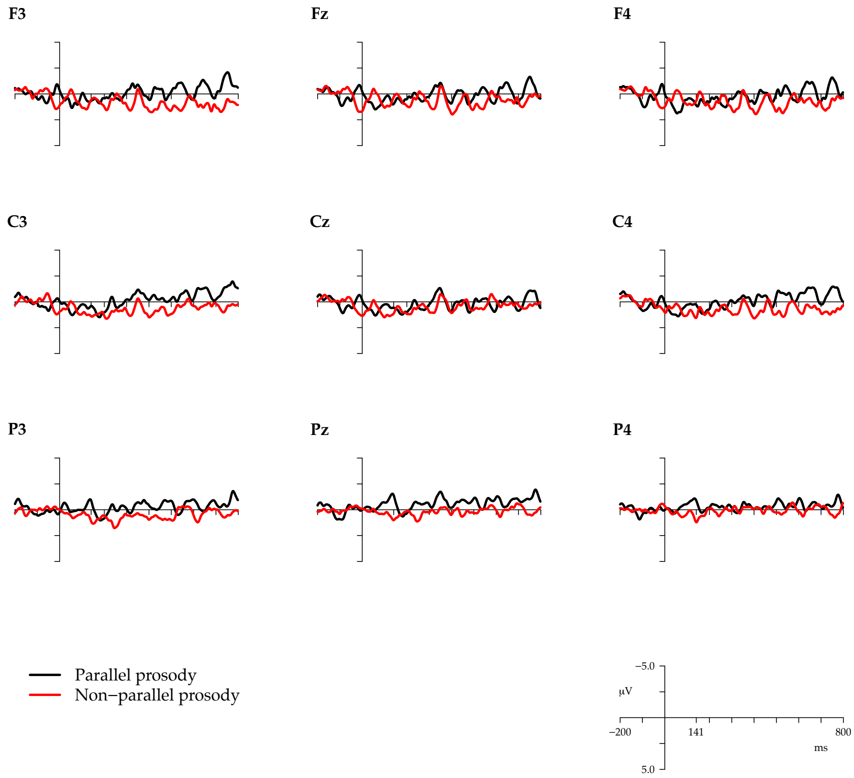


Figure 8.4: Grand averages of Parallel prosody condition (***a***) compared to Non-parallel prosody condition (***b***) at onset (y-axis) of the determiner (*de*) of the object phrase in the second conjunct (*de bloemen*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Approximate onset of the noun (*bloemen*) is 141 ms ($SE = 0.004$) after onset of the determiner. Corresponding example sentences can be found on page 153.

8.5 Discussion

In this experiment I hypothesised that the prosody of the first conjunct predicts the remaining structure in the second conjunct. A LAN-like component between 200-400 ms was expected to be found on the determiner in the Non-parallel prosody condition, reflecting an index of prediction strength. Additionally, I hypothesised that an unexpected noun would elicit an N400 effect, although scepticism was in order given the null effect in the replication study reported in Chapter 5.1. Sentences containing unexpected structures were expected to yield a processing difficulty, as reflected by ERP components, as well

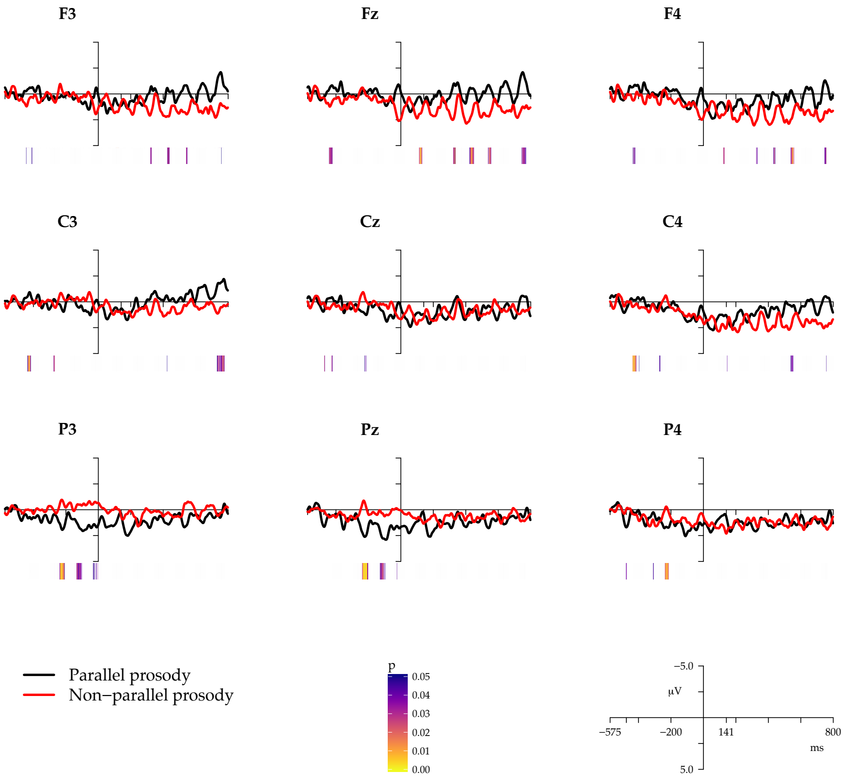


Figure 8.5: Grand averages of Parallel prosody condition (***a***) compared to Non-parallel prosody condition (***b***) at onset (y-axis) of the determiner (*de*) of the object phrase in the second conjunct (*de bloemen*) at electrode sites F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. Baseline of 100 ms starts at -575 ms. Approximate onset of the noun in Object2 (*bloemen*) is 141 ms (SE = 3.74) after onset of the determiner. At each electrode, significant effects ($p \leq .05$) of a permutation test carried out at every sample (by factor the CONDITION) are shown. Corresponding example sentences can be found on page 153.

as by a comprehension scores.

Since the comprehension scores show a ceiling effect, it is difficult to estimate the effect of prosody in terms of behavioural responses. While the questions may have been too easy to answer, it could be that an assumed parallel prosody constraint is lenient enough not to mislead ultimate comprehension. It could also be that such a constraint is not measurable through behavioural means. Despite the ceiling effect, the correlation between the comprehension scores and the working memory test showed the same pattern as in the previous experiments, meaning that low comprehension scores correlate – although not significantly – with low working memory scores.

Although the ERP results are based on a relatively small sample,[1] some effects may be explored and may be corroborated by future research. In this experiment, I have chosen to use naturally recorded stimuli, which means that they were not re-synthesised or edited (except for the onset and offset).

It was hypothesised that modulation of the prosody of the object in the first conjunct would influence the expectations of the processor when encountering the object phrase of the second conjunct. At frontal electrodes, early positive deflections are apparent in the Non-parallel prosody condition starting 100 ms after onset of the determiner of the second object, as well as 100 ms after onset of the object noun. Some ERP researchers assume that P1 components are "wholly or primarily due to the feedforward sweep of activity through the sensory pathway" (Woodman, 2010:2043). With respect to spatial attention on visual processing, a P1 measured at posterior electrodes has been regarded as reflection of increased attention (also sometimes referred to as "selection"; see for a review Hillyard, Vogel, & Luck, 1998) and may be modulated by a participant's state of arousal. This may in turn influence the entire waveform to become more positive beginning with the P1 wave (Luck, 2014:76). A P1 at frontal electrodes has been coined "Frontal Selection Positivity" (FSP) and has also been found mainly in relation to visual stimuli (see for example Michie et al., 1999). In auditory experiments, selective attention has been connected to an early positivity (P50) (Luck, 2014:81). A P50 (peaking between 50-100 ms) may further be modulated by expectancy giving rise to higher amplitudes for unexpected items (Clementz et al., 2002).

Apparently, the deviant prosody condition may have affected the way the processor treats the unexpected object in the second conjunct – giving it more attention. The positive wave may be either extended due to the unexpected item, or – maybe concurrently – higher cognitive components starting at 300 ms reflect processes of reevaluation of the input. Especially at electrode F8 (not visualised), the positivity was significant in time windows of 50 ms and longer, between 300-350 ms, 350-550 ms, 620-670 ms, and 790-880 ms. It should be noted that, apart from attention/selection effects as reflected by an FSP and and a P50, frontal lobe activity has been connected to a wide range of cognitive demands (such as working memory, episodic memory, problem solving, perception), making it difficult to interpret the effects found here as language-specific. However, it may be the case that processes of accommodation, a term used in linguistic theory to refer to mechanisms at the level of discourse and pragmatics (see also Chapter 2.4.1), are connected to these cognitive demands. At best, we could assume that the positivity around 300 ms relates in part to attentive language comprehension – something I have suggested earlier in Chapters 6 and 7.

[1] In contrast to my previous experiments, I included left-handed participants (after visual inspection of grand averages) to arrive at a larger sample; in general, I have come to the conclusion that linguistic research should not be based on data from right-handed people exclusively.

Rather unexpectedly, already at the subject of the second conjunct (starting at approximately 462 ms before the critical measure point) early positive deflections could be seen, mainly at frontal and central electrodes. However, during acoustic analysis of the stimuli it became clear that, in condition ***a***, the means of the fundamental frequency of the subject in the second conjunct differed significantly from the subject in the first conjunct, whereas in condition ***b*** no such difference was apparent. Possibly, the speaker had relatively more difficulty in producing the Non-parallel condition. At the position of subject noun of the second conjunct, only a slight numerical difference between the Parallel and Non-parallel condition was visible. In addition, the pause in condition ***b*** was slightly longer. It is difficult to determine which of the differences could be the origin of the early positivity. Though it is interesting to note that a P50 has been connected to pitch discrimination (Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011).

At central and – more prominently – posterior electrodes a negative deflection was visible starting around 200 ms after onset, which may be related to an N2. In visual experiments, a posterior N2 may reflect some aspect of focusing of attention and categorisation processes of a stimulus. In auditory experiments, the mismatch negativity (MMN) peaks around the same time point, yet, it usually has a fronto-central midline distribution (see also Luck, 2014:85). A tentative conclusion could be that the negativity reflects processing of the difference in pitch and possibly the difference in duration of the preceding pause. Again, I would like to stress that future experiments are needed to corroborate these preliminary findings. For example, stimuli could be synthesised manipulating only the object in the first conjunct, neglecting the apparent natural production differences. A useful method is "cross-splicing". Target items can be constructed by cross-splicing the critical segment from the Parallel condition (i.e. the right conjunct) over the corresponding segment in the Non-parallel condition. In another design, conditions could be included in which the object in the second conjunct *is* deleted. In this case I would predict the reverse pattern, as no effects should be demonstrable if the object in the first conjunct is de-accented (c.f. example (1c) on page 152).

All in all, it seems that prosodic cues – if they are ungrammatical – do not necessarily impact full interpretation. In other words, it seems that a prosodic parallelism constraint violation has relatively minor consequences. These consequences seem to be undetectable in behavioural responses, although their effects may be measurable in terms of brain activity.

8.6 Conclusion

In an auditory ERP experiment, I tested to what extent the prosody of the first conjunct affects expectations of ellipsis in the second conjunct. While a processing difficulty was expected for non-parallel prosody conditions, the manipulation of prosody could not be measured by means of a comprehension test. In addition, hypothesised ERP results could not be established due to an improper baseline. Although the stimuli were recorded by a trained phonetician, it appeared that the critical differences in production of the object in the first conjunct may have impacted the speaker's production of words later in the sentence, leading to an improper baseline. However, and despite the small sample, an exploratory analysis showed ERP effects related to attention/selection processes at the predicted time point. Additional ERP effects were established at time points preceding the object in the second conjunct which may be due to the pitch difference of the subject in the second conjunct as well as the duration of the pause between the conjuncts. Future research is needed to sustain the exploratory findings reported in this chapter that suggest that a prosodic parallelism constraint violation has relatively minor consequences.

CHAPTER 9

Conclusions and future prospects

9.1 Connecting conclusions

In this thesis, I investigated the ellipsis type Gapping and its sub-type Stripping. After an introduction to the topic in Chapter 1, I discussed in Chapter 2 the relevant theoretical background on Gapping where I demonstrated that Gapping has a multidimensional character. Above all, Gapping has been suggested to be a surface anaphor which has led to a focus on the importance of syntactic structure at the ellipsis site. This contrasts with semantic oriented accounts that might consider ellipsis antecedents to be a type of a deep anaphors. Although I started my research by assuming a simple differentiation between structural and non-structural accounts (as is frequently found in reviews on ellipsis), this opposition seems shaky. Even though theoretical approaches may be leaning towards one side (syntactic, semantic) it appears that to successfully account for the distributional properties of Gapping-like constructions, syntactic, semantic and prosodic factors need to be taken into account. A successful account of ellipsis should be able to answer the question what is the proper balance between these factors. Crucially, three closely related questions have been entertained in the ellipsis literature:

- What is the nature of the ellipsis site (i.e. its representation)?
- What is the nature of the antecedent (referred to as the “identity” condition)?
- Under which conditions is ellipsis allowed (referred to as the “licensing” condition)?

Theoretical treatments make no claims about the timing of ellipsis resolution, which makes it difficult to link theory to processing – a topic I touched on in Chapter 4, in which I put forward suggestions to improve this by means of computational linguistic research. However, these questions may be taken up by experimental research. In an attempt to connect theory to experiments, I utilised the mechanism Copy α, which is associated with theories of surface anaphors, and a cue-based mechanism, which relates to theories of deep anaphors. These mechanisms reflect to some extent the divide between syntax-first and constraint-based approaches. Therefore, they were helpful to make hypotheses as to the time course of the recovery of Gapping and Stripping. I proposed a two-stage mechanism based on retrieval and integration processes and proposed that the two mechanisms make different predictions with respect to the time course of ellipsis processing. A copy account may be costly as it comes to retrieval since searching for and finding structure might be more difficult as a function of the size of the structure. Once a fully fledged structure is available, it is expected that integration processes occur with relative ease. Contrastingly, a cue-based account, which is mainly explaining the mechanism of retrieval, predicts the reverse.

Before testing this hypothesis, I replicated an ERP study in Chapter 5 on verb Gapping in Dutch, the results of which pointed to an integration process reflected by late positivities. Making predictions with respect to processes of retrieval and integration and using pretested stimuli that were based on the replicated study I could not corroborate the effect of late positivities in my first ERP experiment on structural complexity as described in Chapter 6, since my proposed measure point appeared to be too late. As a consequence, any effect of retrieval – if present – could not be determined. In a post hoc analysis of critical words earlier in the test sentences, I was able to find preliminary evidence for the start of the retrieval process. However, it was reflected by a positivity rather than an expected ELAN-related ERP component. In a follow-up experiment on Stripping, it appeared that the addition of an adjunct to the deleted structure could modulate both retrieval and integration phases. I argued that an early positivity is sustained by both acoustic, attentional and linguistic cues, possibly directly targeting a semantic representation (or representations), marking the start of the resolution process to retrieve missing information in order to integrate it with the remnant structure. I further suggested that the relative difficulty of integration of retrieved material is then reflected by the secondary positivity which I related to a P600. While the first positivity seems to be an amalgam of different neural generators, it is conceivable that the late positivities are sustained by several integration processes that work in parallel as the distribution of the late positivities could not be connected to one single mechanism.

In Chapter 7, I investigated the impact of semantic complexity in Stripping conditions, again based on stimuli used in the replicated study. Based on theoretical insights, I hypothesised that quantifying expressions may be a burden on mechanisms of movement and/or copying since additional structural information has to be analysed for such mechanisms to work. I extended this postulation to processing, and suggested that a mechanism such as Copy α would predict a structural processing cost during the recovery of the quantifying expressions, which should be reflected as a syntax-related ERP. Since no difference could be established between non-quantifying and quantifying expressions, I concluded that this is problematic for accounts that consider the representation of a possible antecedent for ellipsis as fully fledged syntactic structure. Again, positive deflections were found during the resolution of Stripping, though only for a small group of participants.

In a final experiment reported in Chapter 8, I carried out an auditory ERP experiment to investigate the effect of prosody on the prediction of Gapping constructions, asking to what extent the prosody of the first conjunct predicts upcoming (deleted) structure. For this experiment, I recorded a selection of stimuli that were used in the replicated study. Although the sample was too small to draw clear conclusions, an exploratory analysis pointed to ERP effects related to attention/selection processes at the critical measure point.

In order to to keep participants engaged in the task, I included comprehension questions in all the experiments. Interestingly, I was able to show that these offline data do not always converge with the online EEG data. That is, it is possible to detect extra processing effort that does not appear to impact the comprehension. Furthermore, comprehension scores may deviate from acceptability scores. As I argued in Chapter 4, an understanding of human language benefits from complementary methods, that is, it cannot be based on acceptability scores alone.

I also carried out a working memory test as a means to control for the variation of the capacity of people's working memory systems. In general, the working memory data showed a small, non-significant positive correlation with sentence comprehension scores. In one experiment, there was a large, significant correlation. Based on the differences in ERP data, I concluded that the correlation might have been caused by a difference in attention level. In general, working memory as measured in this study does not seem to play a large role in comprehension of elliptical sentences.

All experimental findings underscored the multidimensional nature of Gapping. In that sense, an answer to the first question listed above cannot be clear-cut since the nature of the ellipsis site appears to consist of different information types. I have argued in favour of the notion of two consecutive processes underlying ellipsis resolution. While in the theoretical literature a distinction has been assumed between identity of the antecedent and the form of the ellipsis site, to my knowledge, it has never been acknowledged that these conceptions may be associated with a processing order. In addition, the psycholinguistic literature has overlooked a possible order of processing steps, which has led to an ambiguous discussion on the question to what extent ellipsis resolution is cost-free. Although a straightforward linkage between theory and processing is problematic, on the basis of the EEG data it seems that the identity of an antecedent corresponds to a representation of information types that are targeted during retrieval processes, while the actual form of the ellipsis site may be understood in terms of the way that these information types are integrated at a secondary stage. In addition, although a licensing constraint regarding prosodic parallelism may be absent in behavioural responses, it can be measured in terms of brain activity.

Despite the fact that theories of ellipsis lack a comprehensive account of timing, meaning it is hard to relate theory to online processes, some theorists do have their doors open (or at least, ajar) to processing data. While they may differ as to the degree of syntactic structure assumed, they all emphasise the importance of semantic representations. As a side effect, they tend to account for ellipsis constructions (of whichever type) using one mechanism. From a processing view, this is preferred since it seems unlikely that every descriptively different ellipsis type is resolved by a uniquely dedicated procedure. In fact, with the current data it appears that ellipsis processing resembles 'normal' sentence processing to a large extent. Sentence comprehension is an incremental process during which incoming information is paired with an interpretation – updating representations step by step. On a word-by-word basis, the processor parses each new incoming word to retrieve the necessary information. Incrementally, the processor postulates phonological, syntactic and semantic representations integrating different information types to construct the meaning of a sentence. Ellipsis resolution differs in terms of the polarity and latency of the ERP component related to the retrieval phase.

9.2 Limitations and future prospects

However interesting the results in this study are, I am confronted with some limitations. As already noted during data analysis, some results are based on explorations of small sample sizes and need to be corroborated by future experiments. Furthermore, such experiments should be done cross-linguistically, using languages other than English, Dutch and Spanish, which are the languages for which experimental data on ellipsis exists. As has become clear, different methods may yield different results, therefore, it is highly recommended that stimuli sets are tested using different methods in order to get a more complete picture. Preferably, analyses of the same data sets should be published concurrently to prevent other researchers from attempting to replicate findings that will never be found. However, this requires patience, which is an underestimated virtue as long as researchers are rewarded on the basis of output.

As mentioned, the driving force behind the current project was the simple differentiation between syntactic and semantic accounts. The initial idea was to link this differentiation to electrophysiological data. Already at the end of the second chapter, I concluded that Gapping cannot be captured in either syntactic or semantic terms as the most promising accounts (will need to) integrate different levels of representation. In an attempt to connect theoretical insights to existing processing accounts, I arrived at a comparison of two behaviourally motivated models of ellipsis processing that are partially reminiscent of the syntax-semantics divide. Despite this, it also became clear that a mapping between existing theoretical insights and processing may not always be straightforward or even justifiable. Therefore, results accumulated in the current study should be interpreted with great caution if one tries to relate them to theory. I have experienced this as a big limitation and I sincerely hope that theoretical and experimental research will begin to reconcile in the near future. For example, to get a better understanding of cues as used in a pointer account, processing research may very well profit from theoretical insights – and vice versa. After all, both approaches aim to investigate one and the same language system. At this point, a particular experience comes to mind: when attending my poster presentation (Ruijgrok, Cremers, Cheng, & Schiller, 2016) during the Ellipsis Across Borders Conference 2016 in Sarajevo, I was very happy to hear Jason Merchant analysing his theoretically-motivated semantic E feature (as proposed in Merchant, 2001) in terms of a experimentally-motivated cue.

Although this research project has concluded, I anticipate embarking on follow-up experiments. The attentive reader may have noticed that one condition of the pretested sentences as reported in Chapter 5.3 have not been tested in an ERP setting, namely the sentence as shown in (1).

(1) Koen verving de kast in de woonkamer, en Judith niet.
Koen replaced the cabinet in the living.room and Judith not
'Koen replaced the cabinet in the hall, and Judith not.'

Comparing *niet* to a control condition that contains *ook* instead will give insight into the way negation is processed. What's more, the stimuli used in the reported (and proposed) ERP experiments could be used in self-paced reading tasks and in eye-tracking experiments, to investigate how a two-stage resolution process can be measured using these techniques. Finally, I would like to get a better understanding of the location of neural generators underlying ellipsis resolution for which I would need to explore these processes using fMRI technique. As usual, in the attempt to find answers, we generate more questions, which I leave to a future me.

Appendix A

Listed below are the sentences used in the pretest reported in section 5.2. The number after a sentence represents the mean rating as obtained in the pretest.

Practice sentences

1. De kikker leest aan de waterkant een boekje over bedreigde filosofen voor aan de reiger. 5.4
2. Wanneer de stratenmaker honger heeft, eet hij een boterham met kaas en een banaan. 6.7
3. In twee uur tijd viel dinsdagnacht bijna veertig centimeter gras en drie liter hagel. 2.85
4. Het is natuurlijk bijzonder vervelend voor de klanten dat zij hinder ondervinden door het uitvallen van het netwerk. 6
5. Het advies aan het kabinet is om geld uit bezuinigingen deels terug te geven aan de burger door middel van lastenverlichtingen. 6.25
6. Hoewel de basketballer heel veel had geoefend en op zijn dieet had gelet, was zijn spelinzicht om te huilen. 6.15
7. Paul gaat trainingen verzorgen op het nationaal trainingscentrum en hij zal ook talenten begeleiden bij onder meer het Europees jeugdkampioenschap. 6.25
8. Sommige automobilisten kunnen zich niet gedragen in de file. 6.6
9. Sabine epileerde de wenkbrauwen van het fotomodel en Evelien stiftte de lippen van de actrice. 6.9
10. Omdat Mariska op de borstel tandpasta deed en Renate op de wasbak, waren de patiënten moe. 2.45
11. Omdat het tuincentrum in de zomer een aanbieding had, verliepen de klanten in de regen. 2
12. Terwijl Maarten een gedicht schreef over de liefde, zaagde Yvonne een ingewikkeld muziekstuk. 3.3

13. Terwijl de koningin de heropening van het monument bewoonde, liet de minister verstek gaan. 4.35
14. Omdat Karolien in verwachting was van een tweeling, zocht ze een extra grote kinderwagen. 6.95
15. Omdat de nieuwslezer veelvuldig hakkelde, liep de uitzending uit. 6.45
16. Terwijl Sjoerd op zijn verjaardag een taart bakte, haalde Gerrit chips en frisdrank. 6.15
17. Terwijl het musje onder de dakpannen een nestje bouwde, pikten de meeuwen aan de vuilniszakken. 6.85
18. Omdat Emma voor het hoofdgerecht de rijst verwarmde en Kevin voor het dessert de peren stoofde, waren de gerechten klaar. 4.75
19. Omdat Jeroen tijdens het hoorcollege de beamer herstelde en Kim tijdens het werkcollege de computer, waren de professoren sprakeloos. 5.4
20. Omdat Daan in het restaurant de drankjes betaalde en Rik in het café, waren de feestgangers beduusd. 5.3
21. Omdat Dennis in de supermarkt een krat retourneerde en Sophie ook, waren de magazijnmedewerkers druk. 5.3

Implausible fillers

P1. Omdat oma de eendjes brood voerde, schonk Jim de zwanen in. 1.95
P2. Omdat Joost op de piano pingelde, tokkelde Cindy op de radiator. 5.2
P3. Omdat Fabian zijn nieuwe schaar testte, sloot zijn vader de lineaal af. 1.55
P4. Omdat Cora het hout in stukken hakte, betwijfelden de anderen het kampvuur. 2.6
P5. Aangezien Els de voorruit van de auto zeemde, vergaf Dina het oliepeil. 1.85
P6. Aangezien Lenie in de Hoofdstraat woonde, kleurde de bloemist in de Langestraat. 2.15
P7. Doordat de atleet vandaag sneller dan vorige week rende, praatten zijn concurrenten slechter. 3.5
P8. Doordat Frederik zijn geweer met kogels laadde, vuurde Gijs zijn pistool af met pindakaas. 2.6
P9. Nadat Esmee de post bij de villa bezorgde, keek de hond luid naar haar. 2.3
P10. Nadat Valerie bij haar moeder langs ging, knipperde Anneke een uur met haar zus. 2.5
P11. Voordat de klas een film over de Burgeroorlog keek, belde de leraar de werkstukken op. 1.8
P12. Voordat de jongens de zaal van het clubhuis verlieten, steeg de loodgieter het lek op. 1.9
P13. Terwijl de kunstenaar dieren in de sneeuw fotografeerde, boetseerde Dionne in het bos computers. 3.9

P14. Terwijl Sebas een verhaal aan Diana vertelde, zwom Marcus een idee aan zijn baas. 1.55
P15. Terwijl Jeffrey in het boek over de politie bladerde, keek Myrthe in het rotsblok. 2.9
P16. Terwijl de werknemers de artikelen in het rek prijsden, kocht de directeur de psychose in. 1.9
P17. Omdat Team 1 de finale van het toernooi bereikte en Team 2 in de derde ronde snurkte, waren de spelers verrukt. 2.6
P18. Omdat de spits de bal in het doel kopte en Geert de paal toeterde, waren de kijkers verward. 2.05
P19. Omdat Jonne een grote kan koffie zette en de gasten de kopjes op de neushoorn stapelden, waren de beveiligers thuis. 1.95
P20. Omdat de DJ een opname van Queen bezat en de dansers een interview van Prince ondervroegen, waren de lezers verstomd. 2
P21. Omdat Fleur in het museum zich verveeld had en Johan in de afvalzak, waren de bezoekers dronken. 1.45
P22. Omdat Marnix de limonade met een rietje dronk en Milou haar koffie met een kruk, waren de aanwezigen treurig. 1.8
P23. Omdat de Noor honderd meter in de zee zwom en de Fillipijn op de stopwatch, waren de vissen weg. 1.45
P24. Omdat de sergeant de soldaten beval op te letten en de luitenant ook, waren de babies binnen. 2.65
P25. Omdat Stan het liefst een groot glas sinas dronk en zijn buurjongen ook, waren de flessen teneergeslagen. 2.35
P26. Omdat Ahmed het antwoord op de prijsvraag wist en zijn broer op de medaille, waren de presentatoren stomverbaasd. 2.4
P27. Omdat Boris bij de pomp benzine tankte en zijn zoon in de fietstas, waren de pomphouders depressief. 1.8
P28. Omdat Mandy het hondje van de buren sloeg en haar moeder de katten van haar zus bespande, waren de mannen verdrietig. 3.2
P29. Omdat Nadja op haar kamer aan de opdracht werkte en haar vriendin op school aan haar suikerspin, waren de leerkrachten eten. 1.35
P30. Omdat de bloemist een touwtje om de tulpen bond en Marit de lelies in een envelop, waren de verkopers opgewekt. 2.05
P31. Omdat de klusjesman een spijker in de muur lachte en Kirsten ook, waren de schilderijen opgehangen. 1.6
P32. Omdat de begeleider het haardvuur kookte en de kinderen ook, waren de huizen warm. 2.15

Plausible fillers

F1. Voordat hij de klantenservice belde, raadpleegde Arno de handleiding. 6.55
F2. Omdat haar broer voor zijn examen was geslaagd, vierde Nandi feest. 6.7

F3. Omdat hij een gebroken been had, ging Olaf niet naar school. 6.75
F4. Voordat hij met het spelletje begon, stelde Omar het niveau in. 6.75
F5. Omdat haar kindje ziek was, kon Eefje niet naar de vergadering. 6.9
F6. Voordat de eerste gasten kwamen, ruimde Cor het huis helemaal op. 6.85
F7. Omdat zijn bedrijf failliet was gegaan, zat de ondernemer in de schulden. 6.75
F8. Nadat de kinderen naar bed waren gegaan, genoot Onno van de rust. 6.95
F9. Omdat de agenten hem hardhandig hadden aangehouden, diende de relschopper een klacht in. 6.8
F10. Omdat de caissière extra snel werkte, hoefden de klanten niet lang te wachten. 6.55
F11. Nadat haar man iets te drinken had ingeschonken, zette Wilma de dvd aan. 6.9
F12. Omdat hij niet goed functioneerde, liet de directeur de medewerker op gesprek komen. 6.25
F13. Nadat Claire in eigen doel had geschoten, gingen de tegenstanders uit hun dak. 6.9
F14. Omdat de oude man haar op de gang aansprak, kwam de verpleegster te laat. 6.4
F15. Omdat Lenny zijn vragen had beantwoord, was de man erg tevreden over de service. 6.05
F16. Voordat de juf er iets van kon zeggen, ruimde Helen haar mobiele telefoon op. 6.15
F17. Omdat de commissie Marjan de beste kandidaat vond, benoemde de voorzitter haar tot secretaris. 6.65
F18. Terwijl Donja veel aandacht aan de lay-out besteedde, richtte Hein zich meer op de inhoud. 6.85
F19. Terwijl Gerda op de bank televisie keek, zat Sanne aan tafel te puzzelen. 7
F20. Terwijl Ron alvast meel en suiker afwoog, zette Laura de keukenmachine klaar. 6.9
F21. Terwijl Rinus zijn dochter een vulpen overhandigde, bood oma het meisje een dagboek aan. 6.2
F22. Terwijl Joke een toneelstuk in de schouwburg bekeek, luisterde haar man naar een concert. 6.55
F23. Terwijl de jongen de slaapzaal van de meisjes veegde, schrobde zijn broer de hal. 6.3
F24. Terwijl Samantha een luchtje van Mexx opspoot, smeerde Tamara zich in met een geurige bodymilk. 6.95
F25. Terwijl Henny naar het centrum fietste, ging haar moeder met de bus naar de buitenwijk. 6.85
F26. Terwijl mijn buurmeisje vertederd naar de zwerfkat keek, lokte de buurman hem naar zich toe. 6.65
F27. Terwijl de brandweer de brand in de woning bluste, hield Ronald de omstanders op afstand. 6.75

F28. Terwijl Harriet voor een gala een jurk paste, schafte Nancy voor een sollicitatie een broekpak aan. 6.85
F29. Terwijl Marcel de ramen van de keuken zeemde, plaatste de glaszetter het raam van de woonkamer. 6.2
F30. Terwijl de gast kritiek op de kwaliteit van het eten uitte, keurde zijn vrouw de wijn af. 6.15
F31. Terwijl de huishoudster de gordijnen spoelde, luchtte Monica de dekbedden. 5.85
F32. Nadat een kwajongen bij de Hema een rookworst stal, vertelde Anke aan de agent hoe hij eruit zag. 5.85
F33. Terwijl Martin de gastvrouw voor de leuke avond bedankte, nodigde de gastheer hem uit nogmaals langs te komen. 5.85
F34. Terwijl de scheidsrechter aan de keeper de bal toekende, warmde de spits langs de zijlijn zijn spieren op. 6.3
F35. Terwijl Pim een studie aan de Hogeschool volgde, studeerde Mathilde aan de universiteit. 6.85
F36. Terwijl Rob de hele wedstrijd voetbalde, wisselde de coach Stefan al na een half uur. 5.9

Test sentences

E1a. Omdat Lisa op de zolder de vaas tekende en Thomas in de tuin de bijen wegjoeg, waren de grootouders tevreden. 6
E1b. Omdat Lisa op de zolder de vaas tekende en Thomas in de tuin de bijen, waren de grootouders tevreden. 5.4
E1c. Omdat Lisa op de zolder de vaas tekende en Thomas in de tuin, waren de grootouders tevreden. 4.8
E1d. Omdat Lisa op de zolder de vaas tekende en Thomas ook, waren de grootouders tevreden. 4.4
E2a. Omdat Jolien voor de repetitie de dans doornam en Richard voor het concert de liedjes neuriede, waren de uitvoeringen succesvol. 4.8
E2b. Omdat Jolien voor de repetitie de dans doornam en Richard voor het concert de liedjes, waren de uitvoeringen succesvol. 5.6
E2c. Omdat Jolien voor de repetitie de dans doornam en Richard voor het concert, waren de uitvoeringen succesvol. 4.8
E2d. Omdat Jolien voor de repetitie de dans doornam en Richard ook, waren de uitvoeringen succesvol. 6
E3a. Omdat Tim in de kamer de vloer sopte en Erik in de bijkeuken de koelkast afnam, waren de huisgenoten verheugd. 5.6
E3b. Omdat Tim in de kamer de vloer sopte en Erik in de bijkeuken de koelkast, waren de huisgenoten verheugd. 6.4
E3c. Omdat Tim in de kamer de vloer sopte en Erik in de bijkeuken, waren de huisgenoten verheugd. 6.2

E3d. Omdat Tim in de kamer de vloer sopte en Erik ook, waren de huisgenoten verheugd. 5.6
E4a. Omdat Julia aan haar moeder de bloemen stuurde en Anouk aan haar vader de kaart schreef, waren de ouders blij. 5.6
E4b. Omdat Julia aan haar moeder de bloemen stuurde en Anouk aan haar vader de kaart, waren de ouders blij. 5.4
E4c. Omdat Julia aan haar moeder de bloemen stuurde en Anouk aan haar vader, waren de ouders blij. 6.8
E4d. Omdat Julia aan haar moeder de bloemen stuurde en Anouk ook, waren de ouders blij. 4.6
E5a. Omdat Robert voor de lunch een quiche serveerde en Inez voor het diner een forel fileerde, waren de gasten opgetogen. 6
E5b. Omdat Robert voor de lunch een quiche serveerde en Inez voor het diner een forel, waren de gasten opgetogen. 5.2
E5c. Omdat Robert voor de lunch een quiche serveerde en Inez voor het diner, waren de gasten opgetogen. 5.6
E5d. Omdat Robert voor de lunch een quiche serveerde en Inez ook, waren de gasten opgetogen. 5.2
E6a. Omdat Britt voor haar dochter een crackertje pakte en Amber voor haar zoon een cola inschonk, waren de kleuters rustig. 6.6
E6b. Omdat Britt voor haar dochter een crackertje pakte en Amber voor haar zoon een cola, waren de kleuters rustig. 5
E6c. Omdat Britt voor haar dochter een crackertje pakte en Amber voor haar zoon, waren de kleuters rustig. 5.4
E6d. Omdat Britt voor haar dochter een crackertje pakte en Amber ook, waren de kleuters rustig. 5.2
E7a. Omdat Mike in het perkje de rozen snoeide en Lotte in het hofje de bomen omhakte, waren de wandelaars teleurgesteld. 5.2
E7b. Omdat Mike in het perkje de rozen snoeide en Lotte in het hofje de bomen, waren de wandelaars teleurgesteld. 3.8
E7c. Omdat Mike in het perkje de rozen snoeide en Lotte in het hofje, waren de wandelaars teleurgesteld. 4.8
E7d. Omdat Mike in het perkje de rozen snoeide en Lotte ook, waren de wandelaars teleurgesteld. 6
E8a. Omdat Lucas in de manege de paarden aaide en Ernst op het erf de geitjes kamde, waren de dieren kalm. 6.2
E8b. Omdat Lucas in de manege de paarden aaide en Ernst op het erf de geitjes, waren de dieren kalm. 6
E8c. Omdat Lucas in de manege de paarden aaide en Ernst op het erf, waren de dieren kalm. 6.2
E8d. Omdat Lucas in de manege de paarden aaide en Ernst ook, waren de dieren kalm. 6
E9a. Omdat Karin bij de kust de bergen schilderde en Michelle op het eiland de bossen doorwandelde, waren de reisgenoten ontspannen. 4.8

E9b. Omdat Karin bij de kust de bergen schilderde en Michelle op het eiland de bossen, waren de reisgenoten ontspannen. 5
E9c. Omdat Karin bij de kust de bergen schilderde en Michelle op het eiland, waren de reisgenoten ontspannen. 4
E9d. Omdat Karin bij de kust de bergen schilderde en Michelle ook, waren de reisgenoten ontspannen. 3.6
E10a. Omdat Niels voor het toetje een likeur meebracht en Femke voor de fondue een kaas smolt, waren de famillieleden voldaan. 6.6
E10b. Omdat Niels voor het toetje een likeur meebracht en Femke voor de fondue een kaas, waren de famillieleden voldaan. 5
E10c. Omdat Niels voor het toetje een likeur meebracht en Femke voor de fondue, waren de famillieleden voldaan. 4.6
E10d. Omdat Niels voor het toetje een likeur meebracht en Femke ook, waren de famillieleden voldaan. 5.6
E11a. Omdat Edwin in de auto de tassen telde en Susan in de bus de kinderen kalmeerde, waren de reizigers opgelucht. 5.2
E11b. Omdat Edwin in de auto de tassen telde en Susan in de bus de kinderen, waren de reizigers opgelucht. 5.6
E11c. Omdat Edwin in de auto de tassen telde en Susan in de bus, waren de reizigers opgelucht. 4.4
E11d. Omdat Edwin in de auto de tassen telde en Susan ook, waren de reizigers opgelucht. 5
E12a. Omdat Robin in het plantsoen de geraniums weghaalde en Milan langs het tuinpad de stenen stapelde, waren de hoveniers ontstemd. 5.8
E12b. Omdat Robin in het plantsoen de geraniums weghaalde en Milan langs het tuinpad de stenen, waren de hoveniers ontstemd. 5.6
E12c. Omdat Robin in het plantsoen de geraniums weghaalde en Milan langs het tuinpad, waren de hoveniers ontstemd. 5.2
E12d. Omdat Robin in het plantsoen de geraniums weghaalde en Milan ook, waren de hoveniers ontstemd. 6.2
E13a. Omdat Hilde in de voortuin het gazon onderhield en Ralph in de achtertuin de paden harkte, waren de buurtgenoten vrolijk. 6.6
E13b. Omdat Hilde in de voortuin het gazon onderhield en Ralph in de achtertuin de paden, waren de buurtgenoten vrolijk. 6.6
E13c. Omdat Hilde in de voortuin het gazon onderhield en Ralph in de achtertuin, waren de buurtgenoten vrolijk. 5.2
E13d. Omdat Hilde in de voortuin het gazon onderhield en Ralph ook, waren de buurtgenoten vrolijk. 5
E14a. Omdat Bianca in het weiland de schapen filmde en Simone op de dijk de tractor fotografeerde, waren de veehouders verrast. 6.6
E14b. Omdat Bianca in het weiland de schapen filmde en Simone op de dijk de tractor, waren de veehouders verrast. 6.2
E14c. Omdat Bianca in het weiland de schapen filmde en Simone op de dijk, waren de veehouders verrast. 5.6

E14d. Omdat Bianca in het weiland de schapen filmde en Simone ook, waren de veehouders verrast. 5.8
E15a. Omdat Maud in dertig minuten de fietswedstrijd aflegde en Mirjam in drie uur de marathon liep, waren de toeschouwers uitgelaten. 6.6
E15b. Omdat Maud in dertig minuten de fietswedstrijd aflegde en Mirjam in drie uur de marathon, waren de toeschouwers uitgelaten. 5.2
E15c. Omdat Maud in dertig minuten de fietswedstrijd aflegde en Mirjam in drie uur, waren de toeschouwers uitgelaten. 5.8
E15d. Omdat Maud in dertig minuten de fietswedstrijd aflegde en Mirjam ook, waren de toeschouwers uitgelaten. 6.2
E16a. Omdat Marloes bij de quiz een prijs bemachtigde en Natalie bij het teamspel een vlag hees, waren de tegenstanders verbitterd. 4.2
E16b. Omdat Marloes bij de quiz een prijs bemachtigde en Natalie bij het teamspel een vlag, waren de tegenstanders verbitterd. 5.6
E16c. Omdat Marloes bij de quiz een prijs bemachtigde en Natalie bij het teamspel, waren de tegenstanders verbitterd. 5.8
E16d. Omdat Marloes bij de quiz een prijs bemachtigde en Natalie ook, waren de tegenstanders verbitterd. 6.4
E17a. Omdat Koen in de woonkamer een kast verving en Judith in de gang een lamp monteerde, waren de bewoners perplex. 6.6
E17b. Omdat Koen in de woonkamer een kast verving en Judith in de gang een lamp, waren de bewoners perplex. 5.4
E17c. Omdat Koen in de woonkamer een kast verving en Judith in de gang, waren de bewoners perplex. 5.8
E17d. Omdat Koen in de woonkamer een kast verving en Judith ook, waren de bewoners perplex. 4.8
E18a. Omdat Laurens na het toernooi de kampioenen prees en Anita tijdens de wedstrijden de sfeer bejubelde, waren de deelnemers ontroerd. 4.4
E18b. Omdat Laurens na het toernooi de kampioenen prees en Anita tijdens de wedstrijden de sfeer, waren de deelnemers ontroerd. 5.2
E18c. Omdat Laurens na het toernooi de kampioenen prees en Anita tijdens de wedstrijden, waren de deelnemers ontroerd. 4.8
E18d. Omdat Laurens na het toernooi de kampioenen prees en Anita ook, waren de deelnemers ontroerd. 5
E19a. Omdat Michiel in de stal de koeien voerde en Lisette in de wei de ganzen observeerde, waren de boeren vergenoegd. 5.4
E19b. Omdat Michiel in de stal de koeien voerde en Lisette in de wei de ganzen, waren de boeren vergenoegd. 4.6
E19c. Omdat Michiel in de stal de koeien voerde en Lisette in de wei, waren de boeren vergenoegd. 4.6
E19d. Omdat Michiel in de stal de koeien voerde en Lisette ook, waren de boeren vergenoegd. 5.6
E20a. Omdat Patrick in het begin de punten maakte en Nicole aan het einde de fout verzweeg, waren de teamgenoten chagrijnig. 4.8

E20b. Omdat Patrick in het begin de punten maakte en Nicole aan het einde de fout, waren de teamgenoten chagrijnig. 4
E20c. Omdat Patrick in het begin de punten maakte en Nicole aan het einde, waren de teamgenoten chagrijnig. 5.4
E20d. Omdat Patrick in het begin de punten maakte en Nicole ook, waren de teamgenoten chagrijnig. 5.2
E21a. Omdat Floor op een zondag de winkel beroofde en David op een vrijdag de vrouw beledigde, waren de inwoners ongerust. 5.2
E21b. Omdat Floor op een zondag de winkel beroofde en David op een vrijdag de vrouw, waren de inwoners ongerust. 5
E21c. Omdat Floor op een zondag de winkel beroofde en David op een vrijdag, waren de inwoners ongerust. 4.2
E21d. Omdat Floor op een zondag de winkel beroofde en David ook, waren de inwoners ongerust. 5.6
E22a. Omdat Marieke voor de party de borden sorteerde en Maaike voor de receptie de glazen vulde, waren de verloofden weltevreden. 5.4
E22b. Omdat Marieke voor de party de borden sorteerde en Maaike voor de receptie de glazen, waren de verloofden weltevreden. 6.2
E22c. Omdat Marieke voor de party de borden sorteerde en Maaike voor de receptie, waren de verloofden weltevreden. 4.4
E22d. Omdat Marieke voor de party de borden sorteerde en Maaike ook, waren de verloofden weltevreden. 5
E23a. Omdat Nina voor het etentje een grill regelde en Ruben voor het buurtfeest een bierpomp aansloot, waren de buren gerustgesteld. 5.8
E23b. Omdat Nina voor het etentje een grill regelde en Ruben voor het buurtfeest een bierpomp, waren de buren gerustgesteld. 5.2
E23c. Omdat Nina voor het etentje een grill regelde en Ruben voor het buurtfeest, waren de buren gerustgesteld. 5.2
E23d. Omdat Nina voor het etentje een grill regelde en Ruben ook, waren de buren gerustgesteld. 5.4
E24a. Omdat Janneke op het terras het hout verplaatste en Sander bij het hek de rotsen opstapelde, waren de boswachters woedend. 6
E24b. Omdat Janneke op het terras het hout verplaatste en Sander bij het hek de rotsen, waren de boswachters woedend. 5.4
E24c. Omdat Janneke op het terras het hout verplaatste en Sander bij het hek, waren de boswachters woedend. 4
E24d. Omdat Janneke op het terras het hout verplaatste en Sander ook, waren de boswachters woedend. 5.6
E25a. Omdat Marijke voor het voorgerecht de ingrediënten klaarzette en Rachel voor de ovenschotel de aardappelen sneed, waren de koks ingenomen. 4.8
E25b. Omdat Marijke voor het voorgerecht de ingrediënten klaarzette en Rachel voor de ovenschotel de aardappelen, waren de koks ingenomen. 3.4
E25c. Omdat Marijke voor het voorgerecht de ingrediënten klaarzette en Rachel voor de ovenschotel, waren de koks ingenomen. 5.8

E25d. Omdat Marijke voor het voorgerecht de ingrediënten klaarzette en Rachel ook, waren de koks ingenomen. 3.8
E26a. Omdat Paula aan de bar de smartlap uitvoerde en Roy op het podium de tango danste, waren de organisatoren goedgehumeurd. 5.6
E26b. Omdat Paula aan de bar de smartlap uitvoerde en Roy op het podium de tango, waren de organisatoren goedgehumeurd. 6
E26c. Omdat Paula aan de bar de smartlap uitvoerde en Roy op het podium, waren de organisatoren goedgehumeurd. 4.4
E26d. Omdat Paula aan de bar de smartlap uitvoerde en Roy ook, waren de organisatoren goedgehumeurd. 4
E27a. Omdat Tessa in de stad de plattegrond vroeg en Simon in de woonwijk de huisnummers wees, waren de adressen gevonden. 3.8
E27b. Omdat Tessa in de stad de plattegrond vroeg en Simon in de woonwijk de huisnummers, waren de adressen gevonden. 5.2
E27c. Omdat Tessa in de stad de plattegrond vroeg en Simon in de woonwijk, waren de adressen gevonden. 3.8
E27d. Omdat Tessa in de stad de plattegrond vroeg en Simon ook, waren de adressen gevonden. 4.4
E28a. Omdat Rosalie tijdens het mentoruur de problemen voorlegde en Guido tijdens de vergadering de oplossingen besprak, waren de leerlingen optimistisch. 4.8
E28b. Omdat Rosalie tijdens het mentoruur de problemen voorlegde en Guido tijdens de vergadering de oplossingen, waren de leerlingen optimistisch. 4
E28c. Omdat Rosalie tijdens het mentoruur de problemen voorlegde en Guido tijdens de vergadering, waren de leerlingen optimistisch. 4
E28d. Omdat Rosalie tijdens het mentoruur de problemen voorlegde en Guido ook, waren de leerlingen optimistisch. 6
E29a. Omdat Fred om half twee de assistent belde en Alwin om half drie de winkeleigenaar mailde, waren de klachten opgelost. 5.6
E29b. Omdat Fred om half twee de assistent belde en Alwin om half drie de winkeleigenaar, waren de klachten opgelost. 5.6
E29c. Omdat Fred om half twee de assistent belde en Alwin om half drie, waren de klachten opgelost. 4.6
E29d. Omdat Fred om half twee de assistent belde en Alwin ook, waren de klachten opgelost. 4.8
E30a. Omdat Anton aan de kassa de brochure afrekende en Jelle aan de balie de tickets annuleerde, waren de medewerkers confuus. 5.4
E30b. Omdat Anton aan de kassa de brochure afrekende en Jelle aan de balie de tickets, waren de medewerkers confuus. 5.4
E30c. Omdat Anton aan de kassa de brochure afrekende en Jelle aan de balie, waren de medewerkers confuus. 5
E30d. Omdat Anton aan de kassa de brochure afrekende en Jelle ook, waren de medewerkers confuus. 4.2
E31a. Omdat Gerrie in het archief de tekeningen kopieerde en Harry in het museum het spijkerschrift vertaalde, waren de conservatoren kribbig. 6.2

E31b. Omdat Gerrie in het archief de tekeningen kopieerde en Harry in het museum het spijkerschrift, waren de conservatoren kribbig. 5.6
E31c. Omdat Gerrie in het archief de tekeningen kopieerde en Harry in het museum, waren de conservatoren kribbig. 5
E31d. Omdat Gerrie in het archief de tekeningen kopieerde en Harry ook, waren de conservatoren kribbig. 5
E32a. Omdat Mira voor het ontbijt de sinaasappels schilde en Erica voor het tussendoortje de courgette meenam, waren de gezinsleden verzadigd. 4.6
E32b. Omdat Mira voor het ontbijt de sinaasappels schilde en Erica voor het tussendoortje de courgette, waren de gezinsleden verzadigd. 4.2
E32c. Omdat Mira voor het ontbijt de sinaasappels schilde en Erica voor het tussendoortje, waren de gezinsleden verzadigd. 5.8
E32d. Omdat Mira voor het ontbijt de sinaasappels schilde en Erica ook, waren de gezinsleden verzadigd. 5
E33a. Omdat Florine in het park de noten verzamelde en Ilse in de boomgaard de appels plukte, waren de picknickmanden vol. 6.2
E33b. Omdat Florine in het park de noten verzamelde en Ilse in de boomgaard de appels, waren de picknickmanden vol. 6.2
E33c. Omdat Florine in het park de noten verzamelde en Ilse in de boomgaard, waren de picknickmanden vol. 5
E33d. Omdat Florine in het park de noten verzamelde en Ilse ook, waren de picknickmanden vol. 4.8
E34a. Omdat Arthur in de fanfare de trompet bespeelde en Isolde in de band de drums verzorgde, waren de festiviteiten geslaagd. 5.6
E34b. Omdat Arthur in de fanfare de trompet bespeelde en Isolde in de band de drums, waren de festiviteiten geslaagd. 5.6
E34c. Omdat Arthur in de fanfare de trompet bespeelde en Isolde in de band, waren de festiviteiten geslaagd. 5
E34d. Omdat Arthur in de fanfare de trompet bespeelde en Isolde ook, waren de festiviteiten geslaagd. 5.6
E35a. Omdat Jennie in de dierentuin de leeuwen bedwong en Leon bij het circus de ezel borstelde, waren de beesten braaf. 5.2
E35b. Omdat Jennie in de dierentuin de leeuwen bedwong en Leon bij het circus de ezel, waren de beesten braaf. 4.4
E35c. Omdat Jennie in de dierentuin de leeuwen bedwong en Leon bij het circus, waren de beesten braaf. 4.4
E35d. Omdat Jennie in de dierentuin de leeuwen bedwong en Leon ook, waren de beesten braaf. 4.8
E36a. Omdat Jos op het dorpsplein de fontein renoveerde en Jet aan de kade de toren besteeg, waren de toeristen verbolgen. 4.4
E36b. Omdat Jos op het dorpsplein de fontein renoveerde en Jet aan de kade de toren, waren de toeristen verbolgen. 4.6
E36c. Omdat Jos op het dorpsplein de fontein renoveerde en Jet aan de kade, waren de toeristen verbolgen. 3.4

E36d. Omdat Jos op het dorpsplein de fontein renoveerde en Jet ook, waren de toeristen verbolgen. 5.2
E37a. Omdat Adrie in de proeverij de wijn goedkeurde en Henk in de fabriek het bier bestelde, waren de producenten gelukkig. 6.2
E37b. Omdat Adrie in de proeverij de wijn goedkeurde en Henk in de fabriek het bier, waren de producenten gelukkig. 6.2
E37c. Omdat Adrie in de proeverij de wijn goedkeurde en Henk in de fabriek, waren de producenten gelukkig. 6
E37d. Omdat Adrie in de proeverij de wijn goedkeurde en Henk ook, waren de producenten gelukkig. 5.6
E38a. Omdat Jan in de winter de tomatenplanten bemestte en Klara in het voorjaar de moestuin inzaaide, waren de opbrengsten geweldig. 5.8
E38b. Omdat Jan in de winter de tomatenplanten bemestte en Klara in het voorjaar de moestuin, waren de opbrengsten geweldig. 5.6
E38c. Omdat Jan in de winter de tomatenplanten bemestte en Klara in het voorjaar, waren de opbrengsten geweldig. 5.6
E38d. Omdat Jan in de winter de tomatenplanten bemestte en Klara ook, waren de opbrengsten geweldig. 4.6
E39a. Omdat Merel in het noorden de ijskap redde en Florentien in het zuiden het regenwoud beschermde, waren de politici overtuigd. 5
E39b. Omdat Merel in het noorden de ijskap redde en Florentien in het zuiden het regenwoud, waren de politici overtuigd. 5
E39c. Omdat Merel in het noorden de ijskap redde en Florentien in het zuiden, waren de politici overtuigd. 3.8
E39d. Omdat Merel in het noorden de ijskap redde en Florentien ook, waren de politici overtuigd. 3.6
E40a. Omdat Arie aan de oostzijde de akkers beheerde en Teun in het westen de vijver dregde, waren de tuinders verbaasd. 4.4
E40b. Omdat Arie aan de oostzijde de akkers beheerde en Teun in het westen de vijver, waren de tuinders verbaasd. 5.2
E40c. Omdat Arie aan de oostzijde de akkers beheerde en Teun in het westen, waren de tuinders verbaasd. 5
E40d. Omdat Arie aan de oostzijde de akkers beheerde en Teun ook, waren de tuinders verbaasd. 5.4
E41a. Omdat Roeland in de schuur een tafel oliede en Mees in de hal een stoel schoonmaakte, waren de huisdieren buiten. 5.6
E41b. Omdat Roeland in de schuur een tafel oliede en Mees in de hal een stoel, waren de huisdieren buiten. 5.2
E41c. Omdat Roeland in de schuur een tafel oliede en Mees in de hal, waren de huisdieren buiten. 5.6
E41d. Omdat Roeland in de schuur een tafel oliede en Mees ook, waren de huisdieren buiten. 4.4
E42a. Omdat Hannes in de avond de autorace finishte en Marleen in de middag de marathon rende, waren de vrienden verbluft. 4.8

E42b. Omdat Hannes in de avond de autorace finishte en Marleen in de middag de marathon, waren de vrienden verbluft. 5.6
E42c. Omdat Hannes in de avond de autorace finishte en Marleen in de middag, waren de vrienden verbluft. 4.2
E42d. Omdat Hannes in de avond de autorace finishte en Marleen ook, waren de vrienden verbluft. 4.4
E43a. Omdat Meike in de slaapkamer de gordijnen ophing en Lizzy in de keuken de wandlamp repareerde, waren de klusjes geklaard. 4.8
E43b. Omdat Meike in de slaapkamer de gordijnen ophing en Lizzy in de keuken de wandlamp, waren de klusjes geklaard. 5.6
E43c. Omdat Meike in de slaapkamer de gordijnen ophing en Lizzy in de keuken, waren de klusjes geklaard. 5.6
E43d. Omdat Meike in de slaapkamer de gordijnen ophing en Lizzy ook, waren de klusjes geklaard. 4.2
E44a. Omdat Lara in de garage een racewagen lakte en Piet op de oprit een spoiler poetste, waren de autoliefhebbers content. 5.4
E44b. Omdat Lara in de garage een racewagen lakte en Piet op de oprit een spoiler, waren de autoliefhebbers content. 5.6
E44c. Omdat Lara in de garage een racewagen lakte en Piet op de oprit, waren de autoliefhebbers content. 4.8
E44d. Omdat Lara in de garage een racewagen lakte en Piet ook, waren de autoliefhebbers content. 5.2

Appendix B

Listed below are the test sentences and fillers used Chapter 7. The number after a sentence represents the mean rating as obtained in the pretest.

Test sentences

E1a. Margot registreerde de namen van de atleten, en Anna ook. 5.1
E1b. Margot registreerde alle namen van de atleten, en Anna ook. 5.4
E2a. Rex sorteerde de foto's van de vakantie, en Lilly ook. 4.9
E2b. Rex sorteerde alle foto's van de vakantie, en Lilly ook. 5.4
E3a. Lara lakte de auto in de garage, en Piet ook. 4.6
E3b. Lara lakte elke auto in de garage, en Piet ook. 5.3
E4a. Max bevestigde de planken aan de wand, en Annelies ook. 5.8
E4b. Max bevestigde alle planken aan de wand, en Annelies ook. 5.9
E5a. Marga navigeerde het vliegtuig naar het zuiden, en Job ook. 4.8
E5b. Marga navigeerde elk vliegtuig naar het zuiden, en Job ook. 5.9
E6a. Dagmar betreurde het rapport van het bestuur, en Florian ook. 6.4
E6b. Dagmar betreurde elk rapport van het bestuur, en Florian ook. 6.3
E7a. Nick voltooide de tocht over de oceaan, en Lars ook. 6.3
E7b. Nick voltooide elke tocht over de oceaan, en Lars ook. 5.6
E8a. Lucas verzorgde de paarden in de stal, en Ernst ook. 5.3
E8b. Lucas verzorgde alle paarden in de stal, en Ernst ook. 4.8
E9a. Karin schilderde de bergen bij de kust, en Michelle ook. 6.0
E9b. Karin schilderde alle bergen bij de kust, en Michelle ook. 5.0
E10a. Melissa serveerde de cake met de nootjes, en Sam ook. 5.3
E10b. Melissa serveerde elke cake met de nootjes, en Sam ook. 4.4
E11a. Tom reinigde de treden van de trap, en Jesse ook. 5.1
E11b. Tom reinigde alle treden van de trap, en Jesse ook. 5.6
E12a. Susan telde de kinderen in de bus, en Edwin ook. 5.6
E12b. Susan telde alle kinderen in de bus, en Edwin ook. 5.3
E13a. Robin zag de geraniums in het plantsoen, en Milan ook. 6.5

E13b. Robin zag alle geraniums in het plantsoen, en Milan ook. 5.5
E14a. Bob stuurde de auto naar het westen, en Victor ook. 4.8
E14b. Bob stuurde elke auto naar het westen, en Victor ook. 5.6
E15a. Bianca filmde de schapen in het weiland, en Simone ook. 5.5
E15b. Bianca filmde alle schapen in het weiland, en Simone ook. 6.0
E16a. Remco kocht de schoenen met de veters, en Hugo ook. 6.1
E16b. Remco kocht alle schoenen met de veters, en Hugo ook. 5.3
E17a. Mirjam volbracht de wandeling in drie uur, en Maud ook. 6.3
E17b. Mirjam volbracht elke wandeling in drie uur, en Maud ook. 5.8
E18a. Ellen noteerde de score van de deelnemers, en Hans ook. 6.3
E18b. Ellen noteerde elke score van de deelnemers, en Hans ook. 6.1
E19a. Natalie veroverde de vlag bij het teamspel, en Marloes ook. 5.1
E19b. Natalie veroverde elke vlag bij het teamspel, en Marloes ook. 5.4
E20a. Tessa vroeg de plattegrond van de stad, en Simon ook. 5.4
E20b. Tessa vroeg elke plattegrond van de stad, en Simon ook. 4.5
E21a. Koen verving de kast in de woonkamer, en Judith ook. 4.8
E21b. Koen verving elke kast in de woonkamer, en Judith ook. 4.6
E22a. Ruud gooide de bal naar het doel, en Kristel ook. 5.5
E22b. Ruud gooide elke bal naar het doel, en Kristel ook. 5.1
E23a. Laurens prees de kampioenen van het toernooi, en Anita ook. 6.1
E23b. Laurens prees alle kampioenen van het toernooi, en Anita ook. 5.8
E24a. Saskia bewonderde de marmot in het hok, en Karlijn ook. 5.5
E24b. Saskia bewonderde elke marmot in het hok, en Karlijn ook. 6.0
E25a. Michiel observeerde de koeien in de stal, en Lisette ook. 6.3
E25b. Michiel observeerde alle koeien in de stal, en Lisette ook. 5.0
E26a. Patrick maakte de punten in het begin, en Nicole ook. 4.6
E26b. Patrick maakte alle punten in het begin, en Nicole ook. 4.5
E27a. Yvonne verzamelde de posters van de zangeres, en Maarten ook. 5.9
E27b. Yvonne verzamelde alle posters van de zangeres, en Maarten ook. 4.8
E28a. Melanie deed de thee in de thermosfles, en Jordy ook. 5.3
E28b. Melanie deed alle thee in de thermosfles, en Jordy ook. 4.6
E29a. Floor beroofde de winkel op een zondag, en David ook. 5.3
E29b. Floor beroofde elke winkel op een zondag, en David ook. 5.1
E30a. Jasper kende de taal van de Kelten, en Gerard ook. 6.1
E30b. Jasper kende elke taal van de Kelten, en Gerard ook. 5.9
E31a. Ben probeerde de spiesjes met de paprika, en Sandra ook. 6.3
E31b. Ben probeerde alle spiesjes met de paprika, en Sandra ook. 5.6
E32a. Richard oefende de liedjes voor het concert, en Jolien ook. 6.3
E32b. Richard oefende alle liedjes voor het concert, en Jolien ook. 5.9
E33a. Alex koerste de waterfiets naar de kade, en Julian ook. 4.9
E33b. Alex koerste elke waterfiets naar de kade, en Julian ook. 5.1
E34a. Rachel sneed de aardappelen voor de ovenschotel, en Marijke ook. 5.3
E34b. Rachel sneed alle aardappelen voor de ovenschotel, en Marijke ook. 4.9
E35a. Rosalie besprak de leerlingen tijdens de vergadering, en Guido ook. 5.1

E35b. Rosalie besprak alle leerlingen tijdens de vergadering, en Guido ook. 5.4
E36a. Jelle betaalde de tickets aan de balie, en Anton ook. 6.0
E36b. Jelle betaalde alle tickets aan de balie, en Anton ook. 6.1
E37a. Mira schilde de sinaasappels voor het ontbijt, en Erica ook. 4.6
E37b. Mira schilde alle sinaasappels voor het ontbijt, en Erica ook. 5.5
E38a. Jet renoveerde de toren aan de kade, en Jos ook. 4.4
E38b. Jet renoveerde elke toren aan de kade, en Jos ook. 5.1
E39a. Carolien bezocht de balzaal van het kasteel, en Mark ook. 5.6
E39b. Carolien bezocht elke balzaal van het kasteel, en Mark ook. 6.1
E40a. Manon legde de hamburger op de barbecue, en Fiona ook. 4.9
E40b. Manon legde elke hamburger op de barbecue, en Fiona ook. 4.6
E41a. Jan bemestte de kamerplant in de winter, en Klara ook. 5.4
E41b. Jan bemestte elke kamerplant in de winter, en Klara ook. 4.6
E42a. Mart likte de lolly van het stokje, en Ida ook. 4.4
E42b. Mart likte elke lolly van het stokje, en Ida ook. 4.5

Related fillers

RF1a. Leo duwde de kano naar de kant, en Marius de auto naar de garage. 4.9
RF2a. Marleen eindigde de marathon in de middag, en Hannes de autorace in de avond. 5.6
RF3a. Margreet gebruikte de verf van de Hema, en Thijme de kwast van de specialist. 5.5
RF4a. Anneke zette de schep in de grond, en Paul de bezem tegen de muur. 5.1
RF5a. Meike hing het gordijn in de slaapkamer, en Lizzy de lamp in de keuken. 4.6
RF6a. Kees vulde de wagen van de boer, en Floris de schuur van de veehouder. 5.9
RF7a. Emma nam de ijscoupe van de dag, en Kevin de koffie van de week. 5.8
RF8a. Femke bracht de kaas voor de fondue, en Niels de likeur voor het toetje. 5.9
RF9a. Paulien verloor het krantje van de sportclub, en Willem de CD van het schoolorkest. 6.1
RF10a. Luuk droeg de mand met het fruit, en Danil de pillen van de apotheek. 5.8
RF11a. Nina regelde de grill voor het etentje, en Ruben de bierpomp voor het buurtfeest. 6.4
RF12a. Janneke verplaatste het hout op het terras, en Sander de rotsen bij de vijver. 6.3
RF13a. Gert transporteerde de hijskraan van de aannemer, en Henri de steiger van de schilder. 6.0

RF14a. Fred belde de medewerkster om twee uur, en Alwin de winkel om half drie. 4.4
RF15a. Sjaak zong het refrein van de zomerhit, en Nico de tekst van de popsong. 5.0
RF16a. Tim sopte de vloer in de kamer, en Erik de koelkast in de keuken. 6.1
RF17a. Henk bottelde het bier in de fabriek, en Adrie de wijn in de proeverij. 6.3
RF18a. Karel opereerde de arm van de gewonde, en Vincent de buik van het slachtoffer. 5.3
RF19a. Florentien beschermde het regenwoud in het zuiden, en Merel de ijskap in het noorden. 5.5
RF20a. Peter ontmaskerde de bedenker van de oplichterij, en Danny de methode van de fraudeur. 5.6
RF21a. Gerrit roemde de knecht vanwege zijn werklust, en Sjoerd de truck vanwege zijn trekkracht. 4.8
RF22b. Marlies verkocht elke hond van de buren, en Irene de sofa van haar ouders. 5.1
RF23b. Betty verwarmde alle aardappelen voor het diner, en Mara de peren voor het dessert. 5.8
RF24b. Mike snoeide alle rozen in het perkje, en Lotte de bomen in de tuin. 5.3
RF25b. Lisa tekende elke vaas voor het raam, en Thomas de bijen in de tuin. 5.1
RF26b. Bram veilde elke CD van de rockster, en Daphne de prenten van de kunstenaar. 5.8
RF27b. Sophie haalde alle planken voor de schutting, en Dennis de kussens voor de tuinstoelen. 5.5
RF28b. Renate bekritiseerde elke staking van de monteurs, en Mariska de houding van de arbeiders. 5.3
RF29b. Wouter repareerde elke band van de fiets, en Chantal de remmen van de brommer. 5.5
RF30b. Jacob waste alle boontjes in wat water, en Rens de borden in een sopje. 5.9
RF31b. Thijs verfde elke deur van de schuur, en Denise de muur in de keuken. 5.0
RF32b. Mieke recenseerde elk boek over de stoornis, en Irma de folder over de ziekte. 5.1
RF33b. Amanda vertroetelde elke kat van de buren, en Josine de kanarie van haar tante. 5.4
RF34b. Micha vervuilde alle bossen van de amazone, en Egbert de rivieren in de jungle. 6.0
RF35b. Stijn gaf alle kaarten van het kwartetspel, en Marie de pionnen van bordspel. 5.6
RF36b. Hidde hechtte elk been van de patiënt, en Agnes de snee van de jongen. 5.0

RF37b. Harry kopieerde elk spijkerschrift in het museum, en Gerrie de tekeningen in het archief. 5.1
RF38b. Ilse raapte alle noten in het park, en Florine de appels in de boomgaard. 5.4
RF39b. Leon temde elk paard op de ranch, en Jennie de leeuwin in de dierentuin. 5.3
RF40b. Emmy vervoerde elke pony van de manege, en Evy de zadels van de ruiters. 5.8
RF41b. Arie beheerde alle akkers aan de oostzijde, en Teun de vijver in het westen. 6.8
RF42b. Roeland oliede elke stoel van de tuinset, en Mees de vloer in de hal. 5.9

Unrelated fillers

UF1. Marcel zeemde de ramen in de keuken, en de glaszetter plaatste het raam in de woonkamer. 5.4
UF2. De huishoudster spoelde de gordijnen in de wasmachine, en Monica luchtte de dekbedden voor het open raam. 5.8
UF3. De scheidsrechter kende de bal toe aan de keeper, en de spits warmde zijn spieren op langs de zijlijn. 5.6
UF4. Ron woog alvast meel, en suiker af, en Laura zette de keukenmachine klaar. 4.9
UF5. Joke bekeek een toneelstuk in de schouwburg, en haar man luisterde naar een concert. 6.0
UF6. Samantha had een luchtje van Mexx, en Tamara smeerde zich in met een geurige bodymilk. 5.8
UF7. Henny fietste naar het centrum, en haar moeder ging met de bus naar de buitenwijk. 6.1
UF8. Mijn buurmeisje keek vertederd naar de zwerfkat, en de buurman lokte hem naar zich toe. 5.8
UF9. Valerie ging langs bij elke grootmoeder, en Anneke bij elke zus. 5.6
UF10. De jongens repareerden elke zaal van het clubhuis, en de loodgieter elk lek. 5.0
UF11. De kunstenaar fotografeerde alle dieren in de sneeuw, en Dionne alle dieren in het bos. 6.1
UF12. Sebas vertelde alle verhalen aan Diana, en Marcus alle sprookjes. 5.0
UF13. Jeffrey bladerde in alle boeken over de politie, en Myrthe in alle kranten. 5.5
UF14. De werknemers prijsden alle snoepjes in het rek, en de directeur alle chips. 5.6
UF15. Team 1 bereikte elke tweede ronde van de toernooien, en Team 2 elke finale. 5.0
UF16. Els zeemde elke voorruit van de auto's, en Dina elke achterruit. 5.1
UF17. Jonne zette alle koffie, en de gasten alle thee. 5.3

UF18. De DJ bezat elke opname van Queen, en de dansers elke plaat van Prince. 5.9
UF19. Fleur verveelde zich in elk museum, en Johan in elke bioscoop. 6.0
UF20. Stan dronk het liefst alle frisdrank, en zijn buurjongen alle koffie. 5.0
UF21. Ahmed wist elke prijsvraag, en zijn broer elke oefenvraag. 5.6
UF22. De klusjesman deed elke spijker in de muur, en Kristen elke schroef. 5.4
UF23. Fabian testte elke nieuwe schaar, en zijn vader elke liniaal. 6.2
UF24. Els zeemde elke voorruit van de auto's, en Dina elke achterruit. 5.9
UF25. Chris vond de nieuwe juf te streng, maar Hannah zei de nieuwe meester heel aardig te vinden. 5.1
UF26. De hovenier bewaterde de rozen met een gieter, maar de teler sproeide de chrysanten met de tuinslang. 5.8
UF27. Petra schaakte sinds haar achtste verjaardag, maar de grootmeester deed al mee aan toernooien op zijn derde. 5.6
UF28. De zakenman bekeek het beursnieuws op internet, maar de ondernemer zwoor bij teletekst. 5.9
UF29. Bert scheerde schapen op de hei, maar Marja coupeerde honden in de villawijk. 5.3
UF30. De acteur droeg zijn prijs op aan zijn ouders, maar Roos bedankte alleen haar vrienden. 6.0
UF31. De raad benoemde de secretaris tot voorzitter, maar de commissie vond Marjan de beste kandidaat. 5.8
UF32. Ivar had een feestmuts op, maar zijn zus een strandhoed. 6.0
UF33. De acteur werd veel geboekt voor commercials, maar het model vooral voor modeshows. 5.7
UF34. Bettina reisde in haar eentje door Azi, maar Frans met een gezelschap door Zuid-Amerika. 5.4
UF35. Ingrid maakte de zalm klaar in de oven, maar Frans de vissticks in de pan. 5.3
UF36. Dieuwertje knuffelde de barbies met de jurkjes, maar de jongen de dinosaurus met de lange staart. 5.0
UF37. De meteoroloog voorspelde regen, maar de waarzegster een onbewolkte toekomst. 4.9
UF38. De elektricien legde de bedrading aan, maar de ICT-er het netwerk. 5.5
UF39. Mariet werkte haar weblog regelmatig bij, maar Odette eens per kwartaal. 5.5
UF40. De docent legde de theorie uit, maar de begeleider het praktijkgedeelte. 5.8
UF41. Rudolf bouwde een hut van planken, maar Pleun een tent van oude gordijnen. 5.3
UF42. Carmen presenteerde een programma op de radio, maar Rein een show op tv. 5.7

Bibliography

Aelbrecht, L. (2007). *Movement and ellipsis: An analysis of gapping.* Paper presented at the 2nd Student Conference on Formal Linguistics, Poznán, Poland.

Aelbrecht, L. (2010). *The Syntactic Licensing of Ellipsis* (Vol. 149). Amsterdam/Philadelphia: John Benjamins.

Arregui, A., Clifton Jr., C., Frazier, L., & Moulton, K. (2006). Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*, *55*(2), 232–246.

Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford: Oxford University Press.

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29.

Baggio, G., Lambalgen, M. v., & Hagoort, P. (2012). Language, Linguistics and Cognition. In R. M. Kempson, T. Fernando, & N. Asher (Eds.), *Philosophy of Linguistics* (pp. 325–355). Oxford: North Holland.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384.

Baltin, M. (2012). Deletion versus pro-forms: an overly simple dichotomy? *Natural Language & Linguistic Theory*, *30*(2), 381–423.

Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language*, 32–68.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and Language Development* (pp. 279–362). New York: Wiley and Sons.

Boersma, P., & Weenink, D. (2017). *Praat: doing phonetics by computer [Computer program].*

Boone, E. (2014). *The syntax and licensing of Gapping and Fragments*. Doctoral Dissertation, Leiden University.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition*, *25*(12), 103–134.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. (2016). A Neurocomputational Model of the N400 and the P600 in Language Processing. *Cognitive Science*.

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Neuroreport*, *18*(17), 1851–1854.

Callahan, S. M., Shapiro, L., & Love, T. (2010). Parallelism Effects and Verb Activation: The Sustained Reactivation Hypothesis. *Journal of psycholinguistic research*, *39*(2), 101–118.

Cambier-Langeveld, T., Nespor, M., & Heuven, V. J. van. (1997). *The domain of final lengthening in production and perception in Dutch.* Paper presented at Fifth European Conference on Speech Communication and Technology, Rhodes, Greece.

Cannon, T. D., Glahn, D. C., Kim, J., Van Erp, T. G. M., Karlsgodt, K., Cohen, M. S., et al. (2005). Dorsolateral prefrontal cortex activity during maintenance and manipulation of information in working memory in patients with schizophrenia. *Archives of General Psychiatry*, *62*(10), 1071–1080.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(01), 77–94.

Carlson, K. (2001). The Effects of Parallelism and Prosody in the Processing of Gapping Structures. *Language and Speech*, *44*(1), 1–26.

Chomsky, N. (1957). *Syntactic Structures*. Berlin: Mouton de Gruyter.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1993). *A minimalist program for linguistic theory. The view from building 20: Essays in linguistics in honor of Sylvain Bromberger, ed. by Kenneth Hale and Samuel Jay Keyser, 1-52*. Cambridge, MA: MIT Press.

Chomsky, N. (2015). Some Core Contested Concepts. *Journal of Psycholinguistic Research*, *44*(1), 91–104.

Chung, S., Ladusaw, W. A., & McCloskey, J. (1995). Sluicing and logical form. *Natural Language Semantics*, *3*(3), 239–282.

Clementz, B. A., Barber, S. K., & Dzau, J. R. (2002). Knowledge of Stimulus Repetition Affects the Magnitude and Spatial Distribution of Low-Frequency Event-Related Brain Potentials. *Audiology and Neurotology*, *7*(5), 303–314.

Clifton. Jr., C. E., & Frazier, L. (2010). Imperfect Ellipsis: Antecedents beyond Syntax? *Syntax*, *13*(4), 279–297.

Coppen, P.-A., Borght van der, Dreumel van, Oltmans, & Teunissen. (1993). Een implementatie van gapping. *Gramma/TTT*, *2*(1), 31–45.

Coppock, E. (2001). Gapping: In defense of deletion. In M. Andronis, C. Ball, H. Elston, & S. Neuvel (Eds.), (pp. 133–148). Chicago: Chicago Linguistics Society.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21–58.

Cremers, C. (1983). On the Form and Interpretation of Ellipsis. In A. ter Meulen (Ed.), *Studies in Modeltheoretic Semantics* (pp. 145–160). Dordtrecht: Foris.

Cremers, C. (1993). *On parsing coordination categorically*. Doctoral Dissertation, Leiden University.

Cremers, C., & Hijzelendoorn, M. (2014). Meaningful grammar is binary, local, anti-symmetric, recursive and incomplete. In J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. O. Schiller, & E. van Zanten (Eds.), *Above and Beyond the Segments: Experimental linguistics and phonetics* (pp. 60–70). Amsterdam: John Benjamins.

Crystal, D. (2010). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

Culicover, P., & Jackendoff, R. (2005). *Simpler syntax*. Oxford: Oxford University Press.

Dalrymple, M., Shieber, S. M., & Pereira, F. C. N. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy*, *14*(4), 399–452.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433.

Dimitrova, D. V. (2012). *Neural correlates of prosody and information structure*. PhD thesis, University of Groningen.

Dobashi, Y. (2009). Multiple spell-out, assembly problem, and syntax-phonology mapping. In J. Grijzenhout & B. Kabak (Eds.), *Phonological Domains: Universals and Deviations* (pp. 195–220). Berlin: Mouton de Gruyter.

Donchin, E. (1981). Surprise! Surprise? *Psychophysiology*, *18*(5), 493–513.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(03), 357–374.

Downie, G., Fay, J., Langlois, J. P., Baker, R., & Sinclair, R. G. (1996). *Gift Shop* [Recorded by The Tragically Hip]. On *Trouble at the Henhouse* [CD]. Bath, Ontario; New Orleans; San Fransisco: MCA.

Drury, J. E., Baum, S. R., Valeriote, H., & Steinhauer, K. (2016). Punctuation and Implicit Prosody in Silent Reading: An ERP Study Investigating English Garden-Path Sentences. *Frontiers in Psychology*, *7*.

Embick, D., & Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: correlational, integrated and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, *30*(4), 357–366.

Fiebach, C., Schlesewsky, M., Lohmann, G., Cramon, D. von, & Friederici, A. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Map-*

ping, *24*(2), 79–91.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic Working Memory and the Establishment of Filler-Gap Dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, *30*(3), 321–338.

Fiengo, R., & May, R. (1994). *Indices and Identity*. Cambridge, MA: MIT Press.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotami (Ed.), *NELS 32* (pp. 113–132). Amherst, MA: GLSA Publications.

Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*. Indiana University Linguistics Club.

Frazier, L. (1987). Syntactic Processing: Evidence from Dutch. *Natural Language & Linguistic Theory*, *5*(4), 519–559.

Frazier, L. (2013). A Recycling Approach to Processing Ellipsis. In L. L.-S. Cheng & N. Corver (Eds.), *Diagnosing Syntax* (pp. 485–501). Oxford: Oxford University Press.

Frazier, L. (2015). Two Interpretive Systems for Natural Language? *Journal of Psycholinguistic Research*, *44*(1), 7–25.

Frazier, L., & Clifton, C. (2005). The syntax-discourse divide: processing ellipsis. *Syntax*, *8*(2), 121–174.

Frazier, L., & Clifton, C., Jr. (2001). Parsing coordinates and ellipsis: Copy α. *Syntax*, *4*(1), 1–22.

Frazier, L., & Clifton, J. (2000). On Bound Variable Interpretations: The LF-Only Hypothesis. *Journal of Psycholinguistic Research*, *29*(2), 125–140.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291–325.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357–1392.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct Neurophysiological Patterns Reflecting Aspects of Syntactic Complexity and Syntactic Repair. *Journal of Psycholinguistic Research*, *31*(1), 45–63.

Garnham, A., & Oakhill, J. (1987). Interpreting elliptical verb phrases. *The Quarterly Journal of Experimental Psychology Section A*, *39*(4), 611–627.

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268.

Ginzburg, J., & Sag, I. (2000). *Interrogative investigations*. Stanford: CSLI publications.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. Y. (2011). Native Experience with a Tone Language Enhances Pitch Discrimination and the Timing of Neural Responses to Pitch Change. *Frontiers in Psychology*, *2*.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Griffiths, J., & Lipták, A. (2014). Contrast and Island Sensitivity in Clausal Ellipsis. *Syntax*, *17*(3), 189–234.

Gussenhoven, C. (1988). Adequacy in intonation analysis: The case of Dutch. In H. Hulst & N. Smith (Eds.), *Autosegmental studies on pitch accent* (pp. 95–121). Dordrecht: Cinnaminson: Foris.

Gussenhoven, C. (2005). Transcription of Dutch Intonation. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 118–145). Oxford: Oxford University Press.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423.

Hale, J. T. (2011). What a Rational Parser Would Do. *Cognitive Science*, *35*(3), 399–443.

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*. Doctoral Dissertation, University of Pennsylvania.

Hartmann, K. (2000). *Right node raising and gapping interface conditions on prosodic deletion*. Philadelphia, PA; Amsterdam: John Benjamins.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1-2), 67–99.

Hijzelendoorn, M., & Cremers, C. (2009). An Object-Oriented and Fast Lexicon for Semantic Generation. *Computation and Language*, *abs/0905.3318*.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353*(1373), 1257-1270.

Hoeksema, J. (2007). Pseudogapping: its syntactic analysis and cumulative effects on its acceptability. *Research on Language and Computation*, *4*(4), 335–352.

Hornstein, N., Nunes, J., & Grohmann, K. K. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jackendoff, R. (2007). Linguistics in Cognitive Science: The state of the art. *The Linguistic Review*, *24*(4), 347–401.

Jackendoff, R. S. (1971). Gapping and Related Rules. *Linguistic Inquiry*, *2*(1), 21–35.

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–

375.

Johnson, K. (2009). Gapping Is Not (VP-) Ellipsis. *Linguistic Inquiry*, *40*(2), 289–328.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kaan, E., Overfelt, C., Tromp, D., & Wijnen, F. (2013). Processing Gapped Verbs. *Journal of Psycholinguistic Research*, 1–32.

Kaan, E., & Swaab, T. Y. (2003). Repair, Revision, and Complexity in Syntactic Analysis: An Electrophysiological Differentiation. *Journal of Cognitive Neuroscience*, *15*(1), 98–110.

Kaan, E., Wijnen, F., & Swaab, T. Y. (2004). Gapping: Electrophysiological evidence for immediate processing of missing verbs in sentence comprehension. *Brain and Language*, *89*(3), 584–592.

Kane, M. J., Conway, A. R. A., & Engle, R. W. (1999). What do working-memory tests really measure? *Behavioral and Brain Sciences*, *22*(1), 101–102.

Kehler, A. (2000). Coherence and the Resolution of Ellipsis. *Linguistics and Philosophy*, *23*(6), 533–575.

Kim, C. S., Kobele, G. M., Runner, J. T., & Hale, J. T. (2011). The Acceptability Cline in VP Ellipsis. *Syntax*, *14*(4), 318–354.

Kobele, G. M. (2015). LF-copying without LF. *Lingua*, *166, Part B*, 236–259.

Kuno, S. (1976). Gapping: A Functional Analysis. *Linguistic Inquiry*, *7*(2), 300–318.

Lappin, S. (1999). An HPSG Account of Antecedent-Contained Ellipsis. In S. Lappin & E. Benmamoun (Eds.), *Fragments: Studies in Ellipsis and Gapping* (pp. 68–97). New York: Oxford University Press.

Lappin, S., & Benmamoun, E. (1999). *Fragments: studies in ellipsis and gapping*. New York [etc.]: Oxford University Press.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*(1), 74–88.

Lawrence, M. (2011). *ez: Easy analysis and visualization of factorial experiments.* (R package version 4.4-0)

Legendre, P., & Legendre, L. (1998). *Numerical Ecology*. Amsterdam: Elsevier.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 94–122). Oxford: Oxford University Press.

Levin, N. S., & Prince, E. F. (1986). Gapping and Causal Implicature. *Papers in Linguistics*, *19*, 351–364.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.

Lewis, S., & Phillips, C. (2015). Aligning Grammatical Theories and Language Processing Models. *Journal of Psycholinguistic Research*, *44*(1), 27–46.

Lobeck, A. (1995). *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford: Oxford University Press.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

MacCoun, R., & Perlmutter, S. (2015). Blind analysis: Hide results to seek the truth. *Nature News*, *526*(7572), 187.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 145–151.

Marr, D. (1982). *Vision*. San Fransisco: Freeman.

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. (2011). Direct-access retrieval during sentence comprehension: Evidence from Sluicing. *Journal of Memory and Language*, *64*(4), 327–343.

Martin, A. E., Nieuwland, M. S., & Carreiras, M. (2012). Event-related brain potentials index cue-based retrieval interference during sentence comprehension. *NeuroImage*, *59*(2), 1859–1869.

McKean, E. (Ed.). (2005). *New Oxford American Dictionary* (Second ed.). New York: Oxford University Press.

Merchant, J. (2001). *The syntax of silence, islands and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, J. (2005). Fragments and ellipsis. *Linguistics and Philosophy*, *27*(6), 661–738.

Merchant, J. (2007). *Three kinds of ellipsis: syntactic, semantic, pragmatic?* Rutgers University.

Merchant, J. (2017). Ellipsis: A survey of analytical approaches. In J. van Craenenbroeck & T. Temmerman (Eds.), *A handbook of ellipsis.* New York: Oxford University Press.

Michie, P. T., Karayanidis, F., Smith, G. L., Barrett, N. A., Large, M. M., O'Sullivan, B. T., et al. (1999). An exploration of varieties of visual attention : ERP findings. *Cognitive Brain Research*, *7*(4), 419–450.

Murphy, G. L. (1985a). Processes of understanding anaphora. *Journal of Memory and Language*, *24*(3), 290–303.

Murphy, G. L. (1985b). Psychological explanations of deep and surface anaphora. *Journal of Pragmatics*, *9*(6), 785–813.

Neijt, A. (1979). *Gapping : a contribution to sentence grammar*. Dordrecht: Foris Publications.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532.

Noort, M. van den, Bosch, P., Haverkort, M., & Hugdahl, K. (2008). A standard computerized version of the Reading Span Test in different languages. *European Journal of Psychological Assessment*, *24*(1), 35–42.

Osterhout, L., & Hagoort, P. (1999). A Superficial Resemblance Does Not Necessarily Mean You Are Part of the Family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 Debate. *Language and Cognitive Processes*, *14*(1), 1–14.

Otten, M., & Van Berkum, J. J. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, *1291*, 92–101.

Peirce, J. W. (2007). PsychoPyPsychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.

Peirce, J. W. (2009). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in Neuroinformatics*, *2*.

Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. J. Reuland & A. ter Meulen (Eds.), *The representation of (in) definiteness* (Vol. 98, pp. 98–129). Cambridge, MA: MIT Press.

Peter, V., McArthur, G., & Crain, S. (2014). Using event-related potentials to measure phrase boundary perception in English. *BMC Neuroscience*, *15*.

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. In W. X. Schneider, A. M. Owen, & J. Duncan (Eds.), *Executive Control and the Frontal Lobe: Current Issues* (pp. 44–54). Berlin: Springer.

Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(3), 878–882.

Phillips, C., & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua*, *151*, 78–95.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Poeppel, D., & Embick, D. (2005). Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 103–120). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Poirier, J., Wolfinger, K., Spellman, L., & Shapiro, L. (2010). The Real-Time Processing of Sluiced Sentences. *Journal of Psycholinguistic Research*, *39*(5), 411–427.

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128–2148.

Pollard, C., & Sag, I. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Pollatsek, A., Treiman, R., & Pollatsek, A. (2015). The Role of Sound in Silent Reading. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading*. New York: Oxford University Press.

R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reckman, H. (2009). *Flat but not shallow: Towards flatter representations in deep semantic parsing for precise and feasible Inferencing*. Doctoral Dissertation, Leiden University.

Reckman, H., Cheng, L.-S., Hijzelendoorn, M., & Sybesma, R. (Eds.). (2017). *Crossroads Semantics: Computation, experiment and grammar*. Amsterdam: John Benjamins.

Repp, S. (2009). *Negation in Gapping*. Oxford: Oxford University Press.

Rietveld, T., & Heuven, V. J. v. (2001). *Algemene fonetiek*. Bussum: Couthinho.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*(1), 75–116.

Ross, J. R. (1967). *Constraints on variables in syntax*. Doctoral Dissertation, MIT.

Ruijgrok, B., Cremers, C., Cheng, L. L.-S., & Schiller, N. O. (2016). *Recovery of elliptical structures: electrophysiological evidence for retrieval and integration processes.* Poster presented at Ellipsis Across Borders 2016, Sarajevo, Bosnia and Herzegovina.

Rumelhart, D., & McClelland, J. (1986). *Parallel Distributed Processing: Foundations*. MIT Press.

Sag, I. (1976). *Deletion and Logical Form*. Doctoral Dissertation, MIT.

Sag, I., & Hankamer, J. (1984). Toward a theory of anaphoric processing. *Linguistics and Philosophy*, *7*(3), 325–345.

Sag, I., & Wasow, T. (2015). Flexible processing and the design of grammar. *Journal of Psycholinguistic Research*, *44*(1), 47–63.

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39.

Sharbrough, F. W., Chatrian, G. E., Lesser, R., Luders, H., Nuwer, M., & Picton, T. (1991). Guideline thirteen: Guidelines for standard electrode position nomenclature american electroencephalographic. *Journal of Clinical Neurophysiology*, 200-202.

Sprouse, J. (2011). A Test of the Cognitive Assumptions of Magnitude Estimation: Commutativity does not Hold for Acceptability Judgments. *Language*, *87*(2), 274–288.

Sprouse, J., & Almeida, D. (2013). The role of experimental syntax in an integrated cognitive science of language. In C. Boeckx & K. K. Grohmann (Eds.), *The Cambridge Handbook of Biolinguistics.* (pp. 181–202). Cambridge: Cambridge University Press.

Staub, A. (2015). Reading Sentences. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading* (pp. 202–216). Oxford: Oxford University Press.

Steedman, M. (1991). Structure and intonation. *Language*, 260–296.

Steedman, M. (1996). *Surface structure and interpretation*. Cambridge, MA: The MIT Press.

Steedman, M. (1999). Connectionist sentence processing in perspective. *Cognitive Science*, *23*(4), 615–634.

Steedman, M. J. (1990). Gapping as constituent coordination. *Linguistics and Philosophy*, *13*(2), 207–263.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142–164.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, *30*(3), 267–295.

Swaab, T. Y., Ledoux, K., Camblin, C., & Boudewyn, M. (2012). Language related ERP components. In S. J. Luck & E. Kappenman (Eds.), *Oxford Handbook of Event-Related Potential Components* (pp. 397–440). New York: Oxford University Press.

Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of Deep and Surface Verb phrase Anaphors. *Language and Cognitive Processes*, *5*(4), 257–280.

Tanenhaus, M. K., Carlson, G. N., & Seidenberg, M. S. (1985). Do listeners compute linguistic representations. *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, 359–408.

Timmer, K., & Schiller, N. O. (2012). The role of orthography and phonology in English: An ERP study on first and second language reading aloud. *Brain Research*, *1483*, 39–53.

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(6), 683–692.

Vicente, L. (2008). *Syntactic isomorphism and non-isomorphism under ellipsis* [Unpublished manuscript]. University of California at Santa Clara.

Wasow, T. (1972). *Anaphoric relations in English*. Doctoral Dissertation, MIT.

Wheeler, R. (2010). *multresp() lmperm.* (R package version 2.1.0)

Williams, E. S. (1977). Discourse and Logical Form. *Linguistic Inquiry*, *8*(1), 101–139.

Winkler, S. (2005). *Ellipsis and Focus in Generative Grammar*. Berlin: De Gruyter Mouton.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031–2046.

Wyngaerd, G. vanden. (2007). *Gapping constituents* [Unpublished manuscript on gapping]. HUB/Crissp/K.U.Leuven.

Yang, G.-Q. R. (2010). Syntactic boundaries and comma placement during silent reading of Chinese text: evidence from eye movements. *Journal of Research in Reading*, *33*(2), 168–177.

Samenvatting in het Nederlands

In deze dissertatie heb ik het type ellipsis “Gapping” (zoals te zien is in (1a) en het sub-type “Stripping” (in (1b)) onderzocht met behulp van de elektrofysiologische methode van “event-related potentials” (ERPs). In de voorbeelden in (1) is *kocht een boek* het antecedent van de ellipsis die in de tweede deelzin plaatsvindt. Kenmerkend voor deze soorten ellipsis is dat er informatie geïnterpreteerd wordt die herwonnen moet worden uit de voorafgaande taalkundige context.

(1) a. De man kocht een boek, en de vrouw een krant. (Gapping)
    b. De man kocht een boek, en de vrouw ook. (Stripping)

De methode van ERPs is bij uitstek geschikt om het tijdsverloop van processen te onderzoeken. Ik heb onderzocht hoe deze types worden verwerkt tijdens woord-voor-woord leestaken en een luistertaak. Wat gebeurt er bijvoorbeeld op het moment dat een lezer (of luisteraar) het woordje *ook* in (1b) verwerkt? Daartoe heb ik één replicatiestudie gedaan en vier zelf ontworpen studies uitgevoerd. Voor die studies zijn ook nog twee “pretests” gedaan om stimuli te testen.

De dissertatie is als volgt opgebouwd. Na een algemene introductie van begrippen en kaders in Hoofdstuk 1, wordt in Hoofdstuk 2 de relevante theoretische literatuur besproken. Het blijkt dat hoewel vaak een differentiatie wordt aangehouden op basis van structurele en niet-structurele benaderingen van Gapping en Stripping, een succesvolle verklaring voor de distributionele aspecten zowel syntactische, als semantische alsook prosodische eigenschappen in ogenschouw genomen moeten worden. Vragen die van belang blijken zijn:

1. In welke termen kan ellipsis worden gerepresenteerd?
2. Hoe kan het antecedent worden gerepresenteerd (wat is de identiteit van het antecedent)?
3. Onder welke condities kan ellipsis plaatsvinden (wat zijn de licenties)?

Theoretische benaderingen maken geen voorspellingen over het tijdsverloop. Dit maakt het moeilijk ze te koppelen aan experimentele data die in Hoofdstuk 3 besproken worden. Ook hierin komen de multidimensionale eigenschappen van ellipsis naar voren. Om toch theorie aan processen te kunnen koppelen wordt in Hoofdstuk 4 voorgesteld dat mechanismen van "Copy α" (Frazier & Clifton, 2001) en "cue-based retrieval" (Martin & McElree, 2008) met elkaar vergeleken kunnen worden. Deze mechanismen weerspiegelen tot op zekere hoogte de differentiatie tussen "syntax-first" en "constrained-based" benaderingen die in Hoofdstuk 2 besproken en vergeleken zijn.

Met behulp van deze mechanismen kunnen voorspellingen worden gedaan over twee stadia waaruit verwerking van Gapping en Stripping bij hypothese bestaat, namelijk, het *vinden* van een mogelijk antecedent (zoals *kocht een boek* in (1)) en het *integreren* van dat antecedent met de overgebleven structuur. Copy α voorspelt dat het vinden van structuur kostbaar is, terwijl juist integratie van gevonden structuur relatief makkelijk is. Cue-based retrieval voorspelt daarentegen het omgekeerde. Er wordt voorgesteld dat dit verschil kan worden gemeten met verschillen in onsets van ERPs. Effecten van vindbaarheid zouden vroeg in het tijdsverloop zichtbaar zijn en effecten van integratie later in het tijdsverloop.

Hoofdstuk 5 rapporteert een replicatie van een studie over de verwerking van Gapping gepubliceerd door Kaan et al. (2013). Net zoals in die studie is een late positieve ERP component gevonden en geïnterpreteerd als effect van een integratieproces. Op basis van de stimuli zoals gebruikt in deze studie, zijn nieuwe stimuli ontworpen en getest om te gebruiken in de voor deze dissertatie ontworpen experimenten die besproken worden in Hoofdstuk 6 en Hoofdstuk 7.

In Hoofdstuk 6 is geprobeerd het effect van de hoeveelheid weggelaten structuur te meten in Gapping en Stripping. Hoewel in het eerste experiment met Gapping zinnen niet altijd een late positieve ERP component kon worden vastgesteld, werd een eerste aanwijzing gevonden voor de start van het vinden-proces. Echter, de component die hierbij gevonden werd (positief van aard met een onset rond de 300 ms) is niet direct te koppelen aan puur syntactisch gerelateerde mechanismen. In een volgend experiment werden stimuli getest zoals in (2). Op de vetgedrukte woorden was telkens het kritische meetmoment. De kleuren van de condities komen overeen met de kleuren die gebruikt zijn in de grafieken in de dissertatie.

(2) a. Omdat Koen een kast verving,
en Judith **ook**, waren de bewoners perplex.
(***VP Stripping***)

b. Omdat Koen een enorme kast verving,
en Judith **ook**, waren de bewoners perplex.
(***VP-Adjective Stripping***)

c. Omdat Koen in de woonkamer een kast verving,
en Judith **ook**, waren de bewoners perplex.
(***VP-Adjunct Stripping***)

d. Omdat Koen in de woonkamer een enorme kast verving,
en Judith **ook**, waren de bewoners perplex.
(***VP-Adjective-Adjunct Stripping***)

Het blijkt dat het toevoegen van een adjunct zoals in condities (2c) en (2d) een effect heeft op zowel het proces van vinden als integreren. Dit is gemeten door twee opeenvolgende positieve componenten op het moment dat de deelnemers het woord *ook* lazen in die condities – vergeleken met conditie (2a).

In Hoofdstuk 7 is onderzocht in hoeverre semantische complexiteit in de vorm van kwantificerende elementen zoals *elke* en *alle* van invloed zijn op het vinden en integreren van weggelaten structuur. Stimuli zoals in (3) werden getest.

(3) a. Mira schilde de sinaasappels voor het ontbijt, en Erica **ook**.
(***Determiner Stripping***)

b. Mira schilde alle sinaasappels voor het ontbijt, en Erica **ook**.
(***Quantifier Stripping***)

De hypothese is dat kwantificerende elementen een tol eisen voor een mechanisme als Copy α. Terwijl in zowel condities met gekwantificeerde geëlideerde elementen (bijvoorbeeld *alle sinaasappels*) als condities met ongekwantificeerde geëlideerde elementen (bijvoorbeeld *de sinaasappels* een tendens van een vroege positieve component werd gevonden, was er geen verschil tussen condities. Met dit experiment wordt weer aangetoond dat het vroege effect niet direct te koppelen is aan puur syntactisch gerelateerde mechanismen.

Hoofdstuk 8 rapporteert een exploratie van een luisterexperiment. Gapping zinnen zoals gebruikt in de studie van Kaan et al. (2013) werden daartoe opgenomen door een fonetisch getrainde spreker. Er is onderzocht in hoeverre de “processor” van de luisteraar op basis van de prosodie van de eerste deelzin (voor de eerste komma) kan voorspellen. In de voorbeeldzinnen zijn woorden die accent dragen met hoofdletters geschreven. Het idee is dat als een object in de linker deelzin accent draagt (*KAART*), de processor een parallele tegenhanger in de rechter deelzin verwacht (*BLOEMEN*).

(4) a. ANOUK zond de KAART aan haar VADER,
en JULIA **de BLOEMEN** aan haar MOEDER.
(***Parallel prosody***)

b. ANOUK zond de kaart aan haar VADER,
en JULIA **de BLOEMEN** aan haar MOEDER.
(***Non-parallel prosody***)

De processor lijkt toch vooral gestuurd te worden door processen die te maken hebben met mechanismen van aandacht en selectie. Tegelijkertijd bleek het niet mogelijk om een effect vast te stellen in de gedragsdata. Beide condities werden even goed begrepen door de deelnemers.

In alle experimenten werd gebruikt gemaakt van een begripstaak. Na elke stimulus-zin volgde een inhoudsvraag over die zin. Het blijkt dat het mogelijk is een verwerkingseffect te meten in termen van ERPs, terwijl dat effect niet zichtbaar is in de begripsdata. Zoals reeds in Hoofdstuk 4 is beargumenteerd, is het inderdaad nodig om gebruik te maken van complementaire methoden van taalkundig onderzoek om tot een theorie over de werking van menselijke taal te komen. Verder is na elk experiment ook een werkgeheugentest afgenomen bij de deelnemers. Vooralsnog lijkt het erop dat werkgeheugen zoals gemeten in deze studie geen grote rol speelt in het begrijpen van elliptische structuren.

In alle experimenten waaraan in totaal 186 mensen hebben deelgenomen, kwam het multidimensionale karakter van Gapping en Stripping naar voren. Een eenduidig antwoord op de eerste vraag zoals hierboven (op pagina 202) gesteld is er dus niet. In de theoretische literatuur wordt een onderscheid gemaakt tussen de identiteit van het antecedent en de vorm (representatie) van de ellipsis. Met deze dissertatie is voor zover we weten voor het eerst getracht deze noties te associëren met een tijdsverloop: de identiteit van het antecedent correspondeert met representaties die tijdens het vinden (stadium 1) van structuur van belang zijn; de vorm van de ellipsis kunnen we begrijpen in termen van de manier waarop deze representaties worden geïtegreerd (stadium 2). Met betrekking tot de derde vraag zoals boven gesteld, is het mogelijk een "licensing constraint" die verband houdt met de prosodie van Gapping constructies te meten in termen van ERPs, terwijl die niet te meten is in termen van gedragsdata.

Hoofdstuk 9 concludeert de dissertatie. Samenvattend verschilt verwerking van ellipsis niet heel veel van de verwerking van 'normale' zinnen. Verwerking van zinnen is een incrementeel proces waarbij inkomende informatie wordt gepaard aan een interpretatie. Bij elke stap worden representaties ververst. Woord voor woord ontleedt de processor de inkomende woorden om de nodige informatie te vinden. Stap voor stap postuleert de processor fonologische, syntactische en semantische representaties die op hun beurt worden geïtegreerd opdat een interpretatie van de zin kan worden vastgesteld. Verwerking van ellipsis verschilt wat betreft ERPs in de polariteit en latentie van de ERP component die te maken heeft met het vinden van het antecedent.

Curriculum vitae

Bobby Ruijgrok was born on 9th December 1972 in Leiden, The Netherlands. Between 1992-2001 he was employed as (assistant) manager in a coffee bar and hi-fi stereo equipment shop. In 2000, he started a furniture design company and he has been co-owner of a music production company since 2008.

In 2006, he started studying General Linguistics at Leiden University taking extracurricular courses at the Leiden University Medical Center and the Psychology department. He obtained his BA degree with distinction in 2010. Following that and also at Leiden University, he enrolled in the MA Clinical and Experimental Linguistics. In 2012, he obtained his MA degree with distinction and he was awarded a grant for a PhD project by the Netherlands Organisation for Scientific Research (NWO) to be carried out at Leiden University Centre for Linguistics (LUCL). In 2015, he was awarded a fellowship for a stay at the Linguistic Summer Institute at the University of Chicago. This dissertation is the culmination of his PhD project.

Since 2010, he has held positions at Leiden University as student assistant, teaching assistant, lab coordinator of the EEG lab, lecturer in methodological courses and lab manager of LUCL's Experimental Linguistics Labs.