THE PSYCHOLOGICAL REALITY OF GRAMMAR

The Theta Principle in Parsing Performance
Published by
LOT
Janskerkhof 13
3512 BL Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6406
email: lot@let.uu.nl
http://www.lotschool.nl

Cover illustration:
‘Garden Path’
Oil on canvas-board
by Oren Sadeh-Leicht

ISBN 978-90-78328-23-0
NUR 632

Copyright © 2007: Oren Sadeh-Leicht. All rights reserved.
THE PSYCHOLOGICAL REALITY OF GRAMMAR

The Theta Principle in Parsing Performance

De Psychologische Realiteit van Grammatica
Het 'Theta Principle' in Taalverwerking
(met een samenvatting in het Nederlands)

Proefschrift
ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. W. H. Gispen,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op vrijdag 23 maart 2007
des middags te 12.45 uur

doork
Oren Sadeh-Leicht
geboren op 7 januari 1973 te Petach-Tikwa, Israël
Promotoren:

Prof. dr. T. M. Reinhart

Prof. dr. E. J. Reuland

Prof. dr. F. N. K. Wijnen
CONTENTS

Acknowledgements ix

1. Foundations: The Thematic Parser 1
   1.1. The question 1
   1.2. ‘Psychological reality’ in competence and performance 4
   1.3. Transparency 7
   1.4. Properties of the thematic parser 9
       1.4.1. The garden path effect and change of thematic domain 9
       1.4.1.1. A false prediction 11
       1.4.2. Principles of the Thematic Transparent Model 13
       1.4.3. The garden path effect as a violation of the OLLC 14
           1.4.3.1. The OLLC saves the false prediction 16
           1.4.3.2. The requirement of immediate reattachment within the OLLC 17
   1.5. Criticisms of the Thematic Transparent Model 22
       1.5.1. Problems with Theta Attachment 22
           1.5.1.1. Head-final languages: ‘immediate attachment’ 23
           1.5.1.2. Evidence in favor of delay 26
           1.5.1.3. Accounting for the putative problems for Theta Attachment 29
           1.5.1.4. Thematic information is not used by the parser in initial resolution of attachments 32
           1.5.1.5. Evidence in favor of use of thematic information 35
       1.5.2. Problems with the OLLC 36
           1.5.2.1. The OLLC is not valid for Japanese 37
           1.5.2.2. The OLLC is not valid for English 38
       1.5.3. Strong transparent approaches to parsing do not correspond with actual performance 40
       1.5.4. Frequency is an alternative account for the thematic parser 41
   1.6. The garden path effect as local ungrammaticality 43
       1.6.1. Internal restructuring & relicensing derive from Adjacency 48
   1.7. Conclusion: a viable model 51

2. Theta Attachment 53
   2.1. Cases of Theta Attachment indeterminability 53
       2.1.1. Cases of theta-role surplus 53
       2.1.2. Non theta-role surplus cases 55
2.2. Predictions of non Theta Attachment models ......................................... 57
  2.2.1. The Garden Path Model .................................................................... 57
  2.2.2. The Diagnosis Model ........................................................................ 59
  2.2.3. The Snip Model .................................................................................. 60
  2.2.4. The LAST Model .............................................................................. 61
  2.2.5. Structural Determinism ..................................................................... 62
  2.2.6. Memory-Based Ranked Parallel Model ........................................... 65

2.3. Two opposing groups of prediction ..................................................... 66

2.4. Preliminary empirical evidence for random performance on theta-role surplus sentences ................................................................. 68

2.5. Experiment 1: Validation of Magnitude Estimation ........................... 70
  2.5.1. Introduction ....................................................................................... 70
  2.5.1.1. Background to the eye movement experiment .......................... 73
  2.5.2. Subjects ........................................................................................... 75
  2.5.3. Materials and design ....................................................................... 76
  2.5.4. Procedure ....................................................................................... 78
  2.5.5. Apparatus ....................................................................................... 78
  2.5.6. Results ............................................................................................. 79
  2.5.6.1. Task comprehension ................................................................. 79
  2.5.6.2. Reply accuracy ............................................................................. 79
  2.5.6.3. Magnitude estimations of complexity ....................................... 79
  2.5.7. Discussion ....................................................................................... 80

2.6. Experiment 2: Magnitude Estimation as a tool for the study of processing difficulty in standard garden paths ................................. 82
  2.6.1. Introduction ....................................................................................... 82
  2.6.2. Subjects ........................................................................................... 82
  2.6.3. Materials and design ....................................................................... 82
  2.6.4. Procedure and apparatus ............................................................... 85
  2.6.5. Results ............................................................................................. 85
  2.6.5.1. Task comprehension ................................................................. 85
  2.6.5.2. Reply accuracy ............................................................................. 85
  2.6.5.3. Magnitude estimations ............................................................... 86
  2.6.5.4. Estimation of occurrence of difficulty ....................................... 87
  2.6.6. Discussion ....................................................................................... 88

2.7. Experiment 3: Theta-role surplus in Dutch ....................................... 89
  2.7.1. Introduction to part A of the experiment ....................................... 89
  2.7.2. Subjects ........................................................................................... 90
  2.7.3. Materials and design ....................................................................... 90
  2.7.4. Plausibility and acceptability pre-tests ........................................... 93
  2.7.5. Procedure and apparatus ............................................................... 93
  2.7.6. Results ........................................................................................... 93
2.7.6.1. Task comprehension .......................................................93
2.7.6.2. Reply accuracy ...........................................................94
2.7.6.3. Magnitude estimations of complexity.............................94
2.7.7. Discussion ........................................................................94
2.7.8. Introduction to part B ..........................................................95
2.7.9. Subjects ...............................................................................95
2.7.10. Materials and design .......................................................95
2.7.11. Procedure and apparatus ..................................................96
2.7.12. Results .............................................................................96
2.7.12.1. Task comprehension .......................................................96
2.7.12.2. Reply accuracy ...........................................................96
2.7.12.3. Magnitude estimation......................................................97
2.7.12.4. Estimations of occurrence of difficulty...........................98
2.7.13. Discussion .........................................................................98
2.8. Conclusion: Theta Attachment as a guiding principle of human
     parsing performance ..............................................................98

3. Center-Embeddings ..................................................................101
3.1. The importance of center-embeddings .....................................101
3.2. Center-embeddings as islands: a new observation ....................102
3.3. Experiment 4: The origin of difficulty of center-embeddings ........105
   3.3.1. Introduction .......................................................................105
   3.3.1.1. Predictions of Rosenbaum and Kim (1976) .................106
   3.3.1.2. Predictions of Miller and Chomsky (1963)..................107
   3.3.1.3. Predictions within the Thematic Transparent Model ....107
   3.3.2. Subjects ..........................................................................107
   3.3.3. Materials and design .......................................................107
   3.3.4. Procedure and apparatus ................................................108
   3.3.5. Results ...........................................................................109
   3.3.5.1. Task comprehension .....................................................109
   3.3.5.2. Magnitude estimations .................................................109
   3.3.6. Discussion ......................................................................110
3.4. Experiment 5: The Similarity-Based and Dependency Locality
     Theories ..................................................................................111
   3.4.1. Introduction to the various predictions ............................111
   3.4.1.1. Predictions of Similarity-Based Interference Theory ....111
   3.4.1.2. Predictions of the Dependency Locality Theory ..........112
   3.4.1.3. Predictions of the Thematic Transparent Model .........113
   3.4.2. Subjects ..........................................................................114
   3.4.3. Materials and design .......................................................114
   3.4.4. Procedure and apparatus ................................................115
Acknowledgements

I have been extremely lucky to work with my three promoters, who hardly subscribe to such a formal definition. I will never be able to surpass Eric Reuland’s immense wisdom, knowledge, and sheer happiness. Frank Wijnen’s kindness, cheerful spirit and knowledge set an example to me. Tanya Reinhart’s sharp, lucid and critical thinking have always had me at awe. But most valuable to me are her genuine care and support over the years. I am privileged to have known them and to have learned from them.

I am grateful to Iris Mulders. Iris has helped a lot to polish parts of this dissertation despite tribulations, provided help with experiments I tried to conduct, and was always willing to lend a hand.

I thank Shalom Zuckerman for welcoming me to his home and family, and for his collegiate support. Our informal chats were important for the experiments conducted here.

My accidental meeting with pleasant Reiko Vermeulen has given further impetus to explore those mysterious center-embedded sentences. Although she did not know me then, Reiko carefully listened to my underdeveloped idea and formalized it. I am grateful for her interest and her Japanese judgments, too. Without her, chapter 3 would have probably never been written.

A most special thank you is due to Tali Siloni for her support and for sparking my interest in syntax. I liked attending Tali’s classes, which so impressed and interested me as an MA student; yet I was never able to imitate her elegant and precise syntactic analyses.

Thanks are due to various faculty members, who happily assisted in practical matters: to Theo Veenker for help in programming, to Huub van den Bergh for the time he invested in showing me all about statistics, and to Herbert Hoijtink for doing the same.

Other researchers have directly or indirectly helped in the shaping process of this dissertation: Jean Aitchison, Sergey Avrutin, Josef Bayer, Irena Botwinik-Rotem, Noam Chomsky, Stephen Crain, Martin Everaert, Janet Dean Fodor, Rachel Giora, Julia Horvath, Mana Kobuchi, Hans van de Koot, Colin Phillips, Hugo Quené, Maaike Schoorlemmer, and Antonella Sorace. Erin O’bryan and Jocelyn Ballantyne have been most helpful in recruiting subjects. I was very lucky to stumble upon (through Tali and Tanya) Bradley Pritchett’s inspiring book and I’d like to extend my admiration to him.

I would like to thank my colleagues/friends for their comments, sentence judgments, discussions, assistance, and the nice times at Lokaal 9: Anna Asbury, Berit Gehrke, Nino Grillo, Çem Keskin, Anna Kijak, Arnout Koornmeef, Shakuntala Mahanta, Øystein Nilsen, and Anca Sevcenco.
At the risk of sounding trite, words are insufficient to express my deepest gratitude to Annemarie Kerkhoff, for absolutely everything. Our encounter was also a strike of good luck. Annemarie has been a true friend one only finds in books, a caring person to rely on, a devoted colleague and a sharp critic. Although we profoundly disagree on many topics related to linguistics, our fervent and long discussions were most fruitful for this dissertation. I do not know whether it would have been completed without her.

My thanks also go to my family for all obvious reasons. This work is dedicated in loving memory of my grandmother Theodora Leicht (1913-2005), who survived Auschwitz by the skin of her teeth.

Oren Sadeh-Leicht

January 7th, 2007
Utrecht, The Netherlands
1. FOUNDATIONS: 
THE THEMATIC PARSER

1.1. The question

The inquiry that follows investigates a question fundamental to the theories of Universal Grammar (UG), human Natural Language Processing (NLP), and human cognition:

(1) Can the study of parsing performance provide evidence that a principle of the theory of Universal Grammar is psychologically real?

From an investigation of the following constructions that lead to processing difficulties, I will be able to argue that their processing difficulty is attributable to a local violation of a single grammatical principle – the Theta Principle (a.k.a. the Theta Criterion):

Garden path
(2) ¿Fat people eat accumulates in the veins.\(^1\)
   (cf. Fat which people eat accumulates in the veins.)

Center-embedding
(3) ¿The rat the cat the dog chased ate died.
   (cf. the dog chased the cat that ate the rat which died).

I will unite the processing difficulties associated with these constructions under a single parsing account tightly related to the theory of Universal Grammar. This will lead to the conclusion that processing difficulty is local grammaticality judgment, a direct reflection of grammar.

\(^1\) The reverse question mark denotes processing difficulty.
Processing difficulties are psychologically real in the sense that they are reflected by elevated neural activity in the human brain. Since I will make a direct association between processing difficulties and the dissatisfaction of a single grammatical principle, the Theta Principle, I will argue that this grammatical principle is psychologically real – represented in the human brain.

I will provide empirical grounding to the theoretical analysis by reporting experiments involving these processing difficulty phenomena. Consequently, the empirical and theoretical studies of parsing performance combined will provide a single answer to question (1):  

(4) The Theta Principle is psychologically real.

In the remainder of this chapter I will introduce the conceptual foundations of the approach to be undertaken here for the inquiry of question (1). I will present the specific model of the human sentence processing mechanism (the parser, henceforth), which must establish permissible relations between already identified constituents. I will employ in the study of the aforementioned processing phenomena the thematic parser, based on Pritchett’s (1992a) seminal work. This chapter further introduces advancements in syntactic theory relevant to this model, especially within the Probe-Goal Framework of the Minimalist Program (Chomsky, 1995, 2005).

I will present objections to the validity of the model made over the years. The discussion of these objections will reveal that the model is nevertheless valid and conceptually correct for both right and left-branching languages. Moreover, the discussion will lead to the conclusion that processing difficulty found in garden path constructions is local ungrammaticality: the parser locally violates a principle of grammar: the Theta Principle.

The difference in use of the term ungrammaticality should be stressed already at this stage. Ungrammaticality is defined generally as a theoretical notion: a sentence not generated by grammar is ungrammatical. However, in the parsing of garden paths, and as I will also demonstrate for center-embeddings, a local violation of a grammatical principle arises. It is in this sense that I use the term (local) ungrammaticality throughout this study.

In the second chapter, I will make specific and unique predictions deduced from the model employed. A series of experiments will be reported designed to test these predictions in relation to prominent psycholinguistic

---

2 Other principles are not excluded, of course, from being psychologically real. It is only that this study will be able to point onto the psychological reality of the Theta Principle.
theories of the garden path phenomenon (the Garden Path Model (Frazier, 1987a), the Diagnosis Model (Fodor and Inoue, 1994), the LAST Model (Townsend and Bever, 2001), the Snip Model (Lewis, 1993), the Memory-Based Ranked Parallel Model (Gibson, 1991), and Structural Determinism (Gorrell, 1995b, Sturt and Crocker, 1996)). The conclusions that I will draw from the experiments and the theory are that the garden path difficulty is attributable to a local violation of a grammatical constraint, namely the Theta Principle, and that only the Thematic Transparent Model is compatible with empirical evidence.

The experiments employ a technique called Magnitude Estimation, which is currently under development for the assessment of subtle differences in grammaticality judgments. Magnitude Estimation is my technique of choice since I regard the garden path phenomenon as a reflection of local ungrammaticality. I adjusted the method to be suitable to answer the experimental questions and will report two validation experiments of the technique. The experiments show that Magnitude Estimation is a reliable technique for the study of processing difficulties.

Chapter 3 expands the model of the thematic parser to center-embedded constructions, and demonstrates that the model is generalizable to phenomena it was originally not meant to explain, providing yet stronger support for its validity.

I will provide the difficulty associated with center-embedded constructions a syntactic account. My assertion relies on a discovery, never noted before as far as I know, that canonical center-embeddings contain embedded complex NP islands, such that the difficulty related to these structures is accounted for by considerations of locality, much like any other complex NP island. In the theoretical investigation of related constructions, I will show that center-embeddings involve a serious local dissatisfaction of the Theta Principle. Unexpectedly, grammar can exclude center-embeddings, in contrast to what has been thought about these structures in the past fifty years. This finding establishes a unified account of seemingly disparate processing phenomena (garden paths and center-embeddings) within a transparent approach.

I will report three experiments that tested claims of various theories about the source of difficulty of center-embeddings (Similarity-Based Interference Theory (Lewis, 1996), Dependency Locality Theory (Gibson, 2000), Miller and Chomsky (1963), and Rosenbaum and Kim (1976)), including predictions of the current model. The experiments point to a clear conclusion: the difficulty associated with center-embeddings can only be explained by a thematic transparent model.

In Chapter 4, I will conclude that the aforementioned constructions, which lead to processing difficulty, are amenable to a single unified account
tightly linked to the theory of Universal Grammar: They all locally violate the Theta Principle, a grammatical principle. In this chapter, I argue for the psychological reality of the Theta Principle as evidenced by parsing performance, and for a universal innate mechanism of human sentence processing.

1.2. ‘Psychological reality’ in competence and performance

It is a common belief that the theory of physics is real: that the constructs of the theory of physics bear a symmetric relation to the laws of the physical world, that indeed the physical world follows the theory of physics. Moreover, it is granted as proof that since the laws of physical theory are confirmed by physical experiments that they are indeed real. Theories of Universal Grammar have been criticized on the basis that they have not been shown to be psychologically real in the sense of physical reality and that they lack both experimental proof and theoretical justification for the principles they incorporate.

This section fleshes out the meaning of the term ‘psychological reality’ within two different fields: grammatical competence and parsing performance. As my main interest is in parsing performance, it seeks to answer the question whether the principles of Universal Grammar are manifested in the temporal (‘incremental’) unfolding of linguistic input. First, the meaning of ‘psychological reality’ is defined within grammatical competence, which does not consider the time course manifestation of linguistic input. Second, it explores the meaning of ‘psychological reality’ as part of the study of parsing performance taking the temporal aspect into perspective.

Whereas theories of Universal Grammar were criticized for not being psychologically real, theories of human Natural Language Processing have been argued to bear on psychological reality. The reason was that they relied on perceptual tests of various kinds (self-paced reading, cross-modal priming, speeded grammaticality judgments, visual-world paradigm, and tracking of eye movements) and recently, neurocognitive applications (Event-Related Potential, functional Magnetic Resonance Imaging, Transcranial Magnetic Stimulation, Positron Emission Tomography, Magnetoencephalography, etc.). There was evidence for psychological reality, and evidence for a grammatical theory. Perhaps this separation appeared to be warranted by Chomsky’s fundamental distinction between competence and performance (Chomsky, 1964), although he himself did not endorse this epistemological separation.
“We will always search for more evidence and for deeper understanding of given evidence which also may lead to change of theory. What the best evidence is depends on the state of the field. The best evidence may be provided by as yet unexplained facts drawn from language being studied, or from similar facts about other languages, or from psycholinguistic experiment, or from clinical studies of language disability, or from neurology, or from innumerable other sources... But there is no distinction of epistemological category.” (Chomsky, 1980, p. 190)

Nonetheless, the epistemological divergence widened from the 1970s on due to the demise of the so-called Derivational Theory of Complexity (DTC). DTC interpreted the version of generative grammar of the time (transformational grammar) as a theory of online linguistic computation, but failed to account for empirical findings supplied by psycholinguistics (Fodor, Bever and Garrett, 1974). More than ever, theories of grammar relied on the examination of languages through informant judgments. Psycholinguistics, on the other hand, turned to language perception experiments. Generally, this situation permeated to present time; the gap widened: the contemporary Minimalist Program seems to have no appeal to most psycholinguistic researchers (Ferreira, 2005), reflecting a general trend.

Chomsky’s opinion is that it is meaningless to ask whether a theory of grammar is real – just like it is meaningless to ask whether physical theory is real, or for that matter, any other theory. Any serious theory of language carries a certain insight:

“Any theory of language, grammar, or whatever carries a truth claim if it is serious, though the supporting argument is, and must be, inconclusive.” (Chomsky, 1980, p. 109)

The only question that could be posed is whether a theory is true, that is, in Chomsky’s denotation of truth: “conforms to external reality” (Chomsky, 1980, p. 27).

The opinion expressed by Chomsky is identical to Albert Einstein’s view of the correspondence between theoretical physics (in the case of language, generative grammar) and the empirical world (say, grammaticality judgments):

---

1 Although Phillips (1996) interestingly argues for the DTC’s validity.
“A complete system of theoretical physics is made up of concepts, fundamental laws which are supposed to be valid for those concepts and conclusions to be reached by logical deduction. It is these conclusions which must correspond with our separate experiences; in any theoretical treatise their logical deduction occupies almost the whole book.” (Einstein, 1933, p. 272)

Psychological reality is a claim to the ‘truth’ of a theory, where ‘truth’ is conformity to or correspondence with external reality. Thus, the theory of UG is psychologically real and the principles it incorporates are ‘true’ because they conform to external reality.

Insofar as this is correct, the principles of UG must be manifested in some way in the processing of linguistic input, too – any sentence in human language would have to eventually comply with the principles of the theory of UG: “both parser and grammar have similar tasks – the association of a structural analysis with a string. How active the competence grammar itself is in the parsing process, however, remains an open question” (Pritchett, 1992a, p. 2).

As stated in the beginning, the main purpose of this inquiry is to provide evidence from the study of parsing performance that a principle of Universal Grammar, the Theta Principle, is psychologically real, i.e. that competence grammar is highly active in the parsing process. Intuitively, the Theta Principle requires that every argument (for the purposes at hand, this is a noun phrase (NP) or prepositional phrase (PP)) be associated with a thematic role, and vice versa, that every thematic role would be associated with an argument (a formal definition follows). As an upshot, this study also remedies a situation reported in the Minimalist Program:

“…what is the status of a violation of the θ-Criterion, whether it involves an unassigned θ-role or an argument lacking a θ-role?
No relevant performance tests are available.” (Chomsky, 1995, p. 314)

The principle can be demonstrated by violating it. In the following sentence,

(5) *Bart killed Homer Lisa.

The verb *kill is transitive, and selects two theta-roles: an external <AGENT> and internal <PATIENT> (or <THEME>) theta-role. The <AGENT> role is given to the argument in the subject position Bart, Homer, an NP argument too, receives the <PATIENT> role in the direct object position. Lisa, an NP argument, requires a theta-role, but non is available. This constitutes a violation of the Theta Principle. In the following sentence,
(6) *The mother neglected.

an argument is missing (What/who did the mother neglect?) for the verb to assign its internal theta-role, and a violation of the Theta Principle excludes this sentence as grammatical.

The violations are termed global, and are normally investigated in analyses of grammatical competence. They are judged ungrammatical (now in the usual meaning: not generated by grammar) and are marked by an asterisk.

Local violations of the Theta Principle are found in parsing performance (as will be shown). They are temporary and reflected by processing difficulty. Hence, processing difficulty is local judgment of grammaticality.

As I will show, the local failure to satisfy the Theta Principle during parsing leads to processing difficulty. To the extent that a local violation of a grammatical principle explains processing difficulty, it is possible to argue that the principle is true (real) by virtue of its conformity (correspondence) to measurable external reality. Analogically, the Theta Principle is psychologically real.

An alternative term for ‘psychological reality’ comes from computational linguistics. There, the ‘goodness of fit’ (equivalent to the conformity to external reality) of a model must be established, in the sense that a certain model must account for a sample or corpus it was designed to account for. In the first case at hand, the sample is garden path constructions and the model is the thematic parser. Chapters 1 and 2 are dedicated to the establishment of the ‘goodness of fit’ of this model in view of criticisms and doubts that have been cast on this type of model over the years.

1.3. Transparency

The preceding section was concerned with the definition of the notion of psychological reality. A more coherent term is the truth of a theory, or ‘goodness of fit’. Nonetheless, I will maintain the use of the term psychological reality in those senses in the continuation.

The argument for psychological reality of a grammatical principle raises the question of the nature of the link between theories of Universal Grammar and phenomena that manifest processing difficulty, or better put, models of parsing performance, which are understood to be two different lines of study. However, they must be linked:

“To the extent that we have an explicit theory of competence, we can attempt to devise performance models to show how this knowledge is put to use.” (Chomsky, 1980, p. 225)
It should be noted that “study of performance relies essentially on advances in understanding of competence” (Chomsky, 1980, p. 226), and that the “study of performance and study of competence are mutually supportive” (Chomsky, 1980, p. 226). How can a link between parsing performance and grammatical competence be ascertained?

Theories of Universal Grammar use the term grammatical competence to denote “a system of rules that generate and relate certain mental representations, including in particular representations of form and meaning” (Chomsky, 1980, p. 90). And in more contemporary terms, derivations, as formulated within the generative research program, are a real property of the brain, “not temporally, but as part of its structural design” (Chomsky, 2000), such that transparent models of parsing performance are the application of grammatical competence in real-time. To be able to make a claim on psychological reality, one must therefore assume strong transparency:

“…the speaker-hearer has internalized a rule system involving the principles of locality and opacity and that judgment and performance are guided by mental computation involving these internally-represented rules and principles.” (Chomsky, 1980, p. 130)

Transparency is crucial to the Minimalist Program, too. The Strong Empirical Hypothesis within the Minimalist Program is that Language should satisfy all empirical tests (Chomsky, 2000). It follows then that the application of grammatical theory (Chomsky, 1995, 1999, 2000, 2005) must stand the empirical test of accounting for processing phenomena, providing additional support to its truth, or ‘goodness of fit’.

The study of phenomena of processing difficulty led Pritchett (1992a) to advocate exactly this position of strong transparency:

“Every principle of the Syntax attempts to be maximally satisfied at every point during parsing.” (p. 138)

The development of a transparent parser goes back a long time and has many versions (e.g. Berwick and Weinberg, 1984, Kimball, 1973, Marcus, 1980, Townsend and Bever, 2001). The opinions on transparency tend to divide into three: there is strong interaction between grammar and parser, a weak interaction or none at all: parser and grammar are one. Strong transparency assumes separation, though interactive, between parser (performance) and grammatical competence.

---

4 As Chomsky (2003) informally puts it: “The language, in that sense, provides instructions to the performance systems” (p. 8). Note that Chomsky regards performance systems as mainly those involved in speech articulation. The transparent parser, however, is a performance system in the sense that it receives input and establishes what the permissible relations between them are as specified by grammar.
grammar (competence). This stance should be distinguished from the weaker claim that linkage between competence and performance is not strong (Boland, 2005a, van de Koot, 1990); or the approach that the parser is grammar (Phillips, 1996).

The relation between grammar and parser plays a crucial role in the decision of the object of study (grammar or parser). For proponents of weak transparency, grammar and parser are two separate fields of study. The parser is a separate module from grammar and they minimally interact (how much interaction there is tends to vary between models). For proponents of the ‘grammar is parser’ position, grammar is all there is, though it is difficult to reconcile the finiteness of the parser with infinite recursion – a fundamental property of grammar.

It should be noted that transparency is unlike the assumption of the Derivational Theory of Complexity (DTC), prevalent in the 1970’s. The DTC surmised that a larger number of transformations should correlate with increase in difficulty. Under the assumption of strong transparency, there is no claim whatsoever as to the correlation between the number of transformations and difficulty.

Be that as it may, it appears that only the approach of strong transparency is able to supply an answer the question (1).

In Section 1.4, I will delineate the properties of a strong transparent parser as suggested by Pritchett. In Section 1.5, I will discuss criticisms of Pritchett’s model to show that they do not necessitate any principle change of the model. In section 1.6, I will argue for the notion that processing difficulty of garden path constructions is local grammaticality judgment.

1.4. Properties of the thematic parser

1.4.1. The garden path effect and change of thematic domain

A person who has a language has access to information about words of the language. The knowledge about words is stored in the lexicon. Of particular interest to Pritchett’s model are words with thematic content. These lexical categories are verbs, nouns and prepositions. The lexical categories specify the ‘thematic structure’: how many arguments these lexical categories license and what semantic roles each argument receives. For example, the verb give must be specified as assigning an <AGENT> role, a <THEME> role, and a <GOAL> role. In (7) John, a book, and Mary have these respective thematic roles.

(7) John gave a book to Mary.

Theta-roles are primary semantic functions of phrases in sentences. The parser’s main goal (of comprehension) is to determine theta relations as soon
as possible (identify theta-roles and assign them). This is the central concern of Pritchett’s model.

To establish thematic roles, the parser follows the principle of Theta Attachment. This principle means that the parser must assign theta-roles as soon as possible and find theta-roles for phrases as soon as possible in a one-to-one mapping between theta-roles and arguments.

Evidently, if the parser commits to a particular theta-role assignment (and therefore satisfies its guiding principle, Theta Attachment), it will be difficult to undo the commitment. The parser may make incorrect assignments, but if the reassigned word does not have to leave its theta domain, then the revision process is easy. A theta domain consists of an element that assigns a theta-role and all of the constituents to which it assigned the theta-role. For example, a prepositional phrase (which is also a syntactic category), consisting of a preposition and a noun phrase, is a theta domain because the preposition assigns a role to the noun phrase (e.g. <THEME>). A verb phrase (a syntactic category, too), consisting of a verb and its complements, is a theta domain as well, because the verb assigns theta-roles of <AGENT>, <THEME>, <GOAL>, <EXPERIENCER>, and so on. We may formally define the notion of theta domain:

(8) **Theta Domain:** $a$ is in the $\gamma$ theta domain of $\beta$ iff $a$ receives the $\gamma$ theta-role from $\beta$ or $a$ is dominated by a constituent that receives the $\gamma$ theta-role from $\beta$.

Tree 1 provides a graphic representation of this simple notion with respect to the verb *kill*.

```
Tree 1
```

For instance, the NP *Mary* is in the <THEME> theta domain of *kill* because it receives the <THEME> theta role from *kill*. It should be emphasized that theta domain is a structural notion, and has no relation to the label of the theta-role. Different verbs have different requirements for theta-roles, and
some verbs are ambiguous, in having more than one set of requirements (the thematic requirements of verbs are referred to as theta-grids). We will deal with these problems later on, but let us see first what revision means.

There are two kinds of revisions. Revisions that maintain theta domains are not costly. Revisions that make a theta-assigned element leave its theta domain are costly. Sentences (9a) and (9b) illustrate the two kinds of revisions and how they depend on the assignment of words to theta-roles.

(9)  a. Without her donations to the charity failed to appear.
     b. Without him donations to the charity failed to appear.
     c. Without her donations to the charity Bob failed to appear.

In (9a) the first assignment concerns her. In its attempt to assign theta-roles as soon as possible, the parser assigns her to the role of <THEME> of without, just as it assigns him to the same role in (9b).

However, when the parser receives donations in (9a) there is no theta-role for it to get if her has the role of <THEME> of without. To provide each word with a theta-role as soon as possible, the parser revises its assignment so that donations becomes the <THEME> of without and her modifies donations. Reassignment of her is not costly, however, since her still falls within the theta domain of without.

In (9a), to the charity is attached as a complement of donations, which retains its assignment as <THEME> of without until failed appears. Failed requires the external theta-role to be assigned, but the only local way to link this role with an argument is to change the assignment of donations from object of without to subject of failed. Since this change involves departure of donations from its current thematic domain, the revision is costly. Costly revision is the garden path effect. In this framework, the term ‘garden path effect’ is reserved for costly reanalysis where readers are misled to commit an element to a certain thematic domain, which turns out to be the incorrect thematic domain.

While (9a) illustrates both costly and uncostly revisions, (9b) has no costly revisions. Sentence (9c), for instance, has cost-free revision of reassigning her from object of without to modifier of donations. Once donations is assigned to object of without, it does not need to be reanalyzed as subject of failed, since Bob is available to occupy that role.

1.4.1.1. A false prediction

Pritchett noted that costly revisions as defined so far (leaving a thematic domain) may make a false prediction with regard to the occurrence of costly revision. Consider the following sentence and its parsing steps:
(10) They gave her photos to Doug.

a. **They gave her**: First, *they* is identified as a pronoun and projects. Since no thematic information is available, the parser proceeds. The next element is *gave*. It is identified as a verb that can assign theta-roles.

*They* is licensed by the <AGENT> theta-role. The argument noun phrase *her* immediately following the verb receives the <GOAL> role of the double object verb: \[TP \{NP They\} \{VP [v gave] \{NP her\}\}\].

b. **photos**: The element is identified as an NP argument. It is licensed by the <THEME> role. The current structure is thus: \[TP \{NP They\} \{VP [v gave] \{NP her\} \{NP photos\}\}\] and not \[TP \{NP They\} \{VP [v gave] \{NP her photos\}\}\]. Why is the former structure chosen, whereas in (9a), it was the latter (i.e., in the latter example *her* was identified as a possessor of *donations*)? This is due to the parser’s ambition to have one-to-one mapping between theta-roles and arguments. In (9a), when *donations* is encountered, the only available role is the preposition’s <THEME> role, which was given to *her*. This means that the parser prefers associating *donations* with the prepositional thematic domain over having it stranded without a theta-role. The way it can accomplish this is by reanalyzing *her* as a possessor. In (10), however, there is a thematic role available for *photos* from the ditransitive verb *give* (<THEME>), which is readily assigned to *photos*.

c. **to Doug**: This elements is identified as the true <GOAL>, which forces a change in thematic domains. Recall that a thematic domain is construed as a structural position, and does not relate to role label. Now, *her* is reinterpreted as possessor, and remains in its original theta domain/position. The element *photos* itself must be reinterpreted as the head of the first rather than the second internal argument, which is erroneously predicted to be costly (following Pritchett’s (1992a) tertiary notation, p. 99):

```
Tree 2
```

```
```

```
```
To correct the false prediction, Pritchett suggested a purely structural constraint on possible revisions, the On-Line Locality Constraint (OLLC). In §1.4.2, I will first make explicit the relevant parsing principles of Pritchett’s model and demonstrate the parsing steps on a garden path sentence in §1.4.3. After the principles of the model and its workings are spelled out, I will explain how the model correctly accounts for sentence (10).

1.4.2. Principles of the Thematic Transparent Model

As we have seen, the parser must establish the ‘argument structure’ of the input during parsing. The guiding principle behind this is Theta Attachment:

(11) **Theta Attachment:** The Theta Principle attempts to be satisfied at every point during processing given the maximal theta grid. (Adapted from Pritchett, 1992a, p. 68)

Theta Attachment by itself must satisfy Chomsky’s (2000) Theta Principle, a principle of grammar. As part of Theta Attachment, the Theta Principle is defined as:

(12) **The Theta Principle:** Pure Merge in theta position is required of (and restricted to) arguments.

The Theta Principle intuitively expresses the need for every argument NP/PP to be licensed by a theta-role, and that every theta-role needs to be awarded to an argument NP/PP (formal definitions withstanding). The Theta Principle is “…implicit in the conception of theta-roles as a relation between two syntactic objects, a configuration, and an expression selected by its head” (Chomsky, 2000, p. 103).

When, for example, a verb is identified during parsing, and its theta-roles are retrieved from the lexicon, the theta-roles also license ‘thematic positions’. Arguments undergo Pure Merge with these thematic positions; and Pure Merge is Merge without moving any element. In the rest of this study, I will use pMerge for Pure Merge.

---

5 Theta Attachment is a direct upshot of the Probe-Goal framework within the Minimalist Program. To see this, it is useful to use the notion of phase (Chomsky, 1999, 2005). Phases are syntactic objects (phrases): CP and v*P. The functional head of the v*P phase is associated with full argument structure, transitive and experiencer constructions. Satisfying Theta Attachment thus means reaching a full proposition, a v*P phase. Theta Attachment may be defined, then, as ‘attempt to complete a v*P phase’.

6 The definition of the Theta Principle to be adopted here differs only terminologically from the one used by Pritchett, which was originally borrowed from Government and Binding (GB) theory.

7 Pure Merge: Merge that is not part of Move (Chomsky, 2000). Merge takes a pair of syntactic objects (SO_i, SO_j) and replaces them by a new combined syntactic object SO_{ij}. 

The structural constraint on revision is the OLLC, as mentioned:

(13) **On-Line Locality Constraint (OLLC):** The target position (if any) assumed by a constituent must be governed by its source position (if any), otherwise attachment is impossible for the automatic Human Sentence Processor. (Pritchett, 1992a, p. 101)

1.4.3. The garden path effect as a violation of the OLLC

In this subsection, I will demonstrate the parsing steps on an example that has both costly and uncostly revisions and will show how these revisions are accounted for by the OLLC. Consider the course of the parse of (14) in accordance with the given principles:

(14) ¿After Fred eats food disappears.

a. **After:** The element is identified as a preposition, with a single theta-role in its maximal theta-grid. Since no potential target for Theta Attachment is available, the role cannot be discharged. At first sight, this may seem like a local violation of Theta Attachment (there is a theta-role, but no argument), but Pritchett points out that:

“…temporary unavoidable violations of parsing principles which arise because of the inherently incremental nature of natural language processing cannot be considered significantly costly.” (Pritchett, 1992a, p. 14)

b. **Fred:** This element is identified as an NP and a potential argument. Since a theta-role and a nominal target are now available, the NP can pMerge with the thematic position licensed by the PP in local satisfaction of Theta Attachment.

---

8 Government: \( \alpha \) governs \( \beta \) iff \( \alpha \) m-commands \( \beta \) and every \( \gamma \) dominating \( \beta \) dominates \( \alpha \). A maximal projection. The definition of m-command is c-command with ‘branching node’ replaced by maximal projection; m-command: \( \alpha \) m-commands \( \beta \) iff \( \alpha \) does not dominate \( \beta \) and every \( \gamma \) that dominates \( \alpha \) dominates \( \beta \), a maximal projection (taken from Pritchett, 1992a, p. 173, n. 101).

9 For ease of exposition, only government will be presented, although the OLLC relates to domination, too.

10 Following Pritchett, the actual label of the role is not significant, only its structural implications are. Labels are utilized here for expository reasons.

11 The parsing algorithm can be informally summarized: (a) Input a word; (b) Recover lexical information, including category and theta grid, and project the appropriate XP(s); (c) Maximally satisfy the Theta Principle as constrained by the OLLC; (d) Affirm that the resulting structure satisfies all relevant grammatical principles (success); and if not (failure) invoke conscious reanalysis, by definition yielding the garden path effect; otherwise continue to the next word (adapted from Pritchett, 1992a, p. 96).
c. **eats**: The constituent is identified as a verb, and its maximal theta-grid is retrieved from the lexicon, containing an `<AGENT>` role, and a `<THEME>` role.

At this stage, the parser realizes that the thematic position licensed by the `<AGENT>` role of *eat* has no local argument with which it may pMerge. Better satisfaction of the Theta Principle entails the move of the NP *Fred* from its current thematic domain to the thematic domain of the verb.

Structurally, the NP *Fred* is initially found in the complement position of the PP. This is its source position. The NP *Fred* must move to the subject position of *eat*, which is its target position. This move is represented in Tree 3 below. The source position of the NP *Fred* is the direct object position of the PP, marked by a double-lined box. Its target position is the subject position [Spec, TP], indicated by a shadowed double-lined box.

Since the target position is governed by the source position, revision is easy: it does not violate the OLLC. This type of reanalysis is therefore permitted.

d. **food**: The element is identified as an NP and a potential argument. In accordance with Theta Attachment and in satisfaction of the Theta Principle, it pMerges with the object position of *eat* licensed by the `<THEME>` thematic role.

e. **disappears**: This element enters the parse, and is identified as a verb. It has a single theta-role in its maximal theta grid. The verb forces revision to allow the discharge of its `<THEME>` role (realized on the subject position). However, this requires the detachment of the NP *food* from its current position and raising it to the subject position of *disappear*, as shown in Tree 4.
The source position of the NP (the direct object position of *eat*) does not govern the target position (subject position of *disappear*). The required reanalysis is impossible for the parser; it violates the OLLC. The violation of the OLLC invokes processing difficulty.\(^{12}\)

1.4.3.1. The OLLC saves the false prediction

I return now to the false prediction found in (10), repeated as (15), and rectify it with the OLLC.

(15) They gave her photos to Doug.

The parse is conducted according to Theta Attachment as explained in §1.4.2. Just before the appearance of *to Doug*, this is the structure so far established (Tree 5):\(^{13}\)

\(^{12}\) As a side note, it may first appear that readers experience difficulty when the final verb is introduced. However, there is no direct prediction of the particular time readers notice difficulty, if at all. The parser appears, then, to be a paradigmatic example of a modular input system in the sense of J. A. Fodor (1983).

\(^{13}\) Following Pritchett’s notation.
With the entrance of *to Doug* to the parse, *photos* must leave its current thematic position. This step is depicted in Tree 6.

Recall that leaving a thematic domain was supposed to be costly. However, sentence (15) does not cause conscious processing difficulty. Examining the OLLC, the source position of *photos* (the second argument position; note that what matters are thematic positions, not thematic labels) governs the target head position, N’, since it governs N’s maximal NP projection. The OLLC therefore remedies the false prediction relying on thematic domains and serves to show that costly reanalysis is structure-dependent.

1.4.3.2. The requirement of immediate reattachment within the OLLC

Some sentences, which appear to have the same structures of garden path sentences, do not actually yield processing difficulty. Consider the following pair; the first incurs a garden path effect, the second does not.
Sentence (16) from Bever (1970) poses one of the most severe processing difficulties known, so much so that it has become a canonical example of a garden path. It contains main clause-relative NP ambiguity. Sentence (17) resembles (16), but is not a garden path. How can this contrast be explained within the current model?

Let us first establish why sentence (16) invokes processing difficulty. We must look for a violation of the OLLC.

a. The horse: this element is identified as a DP. No thematic attachments can be done in the absence of a theta assigner.

b. raced: the constituent is identified to be locally ambiguous between a transitive verb in the past tense, passive participle and past intransitive. Which option should the parser adopt?

(i) Should the parser choose the transitive entry, only the <AGENT> theta-role would be locally discharged (see Tree 7), but not the <THEME> role.

(ii) If the parser were to choose the passive interpretation, then the <THEME> role would be discharged on an empty category (a gap) indexed with a (null) operator, in a relative construction (Tree 8).
Although in this case all theta-roles of the passive verb are discharged, the analysis additionally constructs a higher DP (the relativized NP itself headed by *horse*), for which no role is locally available.

Although a single thematic role is discharged in the passive and transitive analyses; in the transitive analysis the <THEME> awaits discharge and in the passive analysis there is a DP awaiting a theta-role. The past intransitive analysis must be pursued, simply because it allows immediate satisfaction of the Theta Principle.

c. *past the barn*: the constituent is identified and a PP is projected. It is adjoined, receiving the (optional) directional <GOAL> role of *race*.

d. *fell*: the verb’s <THEME> role is recovered. There is no local argument that can be pMerged with the thematic position licensed by the <THEME> role of *fall*.

The Theta Principle can be satisfied by relicensing the *horse raced past the barn* in the past intransitive interpretation as a relative NP, thus allowing *fall* to pMerge its <THEME> thematic position with an argument.

This requires *horse*, which originally pMerged with the thematic position licensed by the <AGENT> role of *race*, to be reanalyzed as the (head of the) subject of *fall*. The relative NP pMerges with the thematic position licensed by the <THEME> role of *fall*. The resulting final structure is found in Tree 9.
The reanalysis involves raising from one subject position to another. Since the source position of the horse does not govern its target position, the OLLC is violated, and processing difficulty ensues.

Prima facie, sentence (17) should also yield processing difficulty. When disintegrate is encountered, the spaceship must be reanalyzed from the subject position of destroy, to the subject position of disintegrate, just like horse in sentence (16). A garden path effect would be erroneously predicted.

Notice first that in this case an intransitive entry is unavailable for destroy. Suppose the parser selects the transitive entry. Pritchett (1992a) notes that the parser, “employing grammatical knowledge directly” (p. 92), realizes that by the time it receives the PP in the battle, a direct object is missing: “a nominal complement must occur right-adjacent to its verb, unseparated from it by other arguments or adjuncts” (p. 94). The absence of an obligatory internal argument right-adjacent to the verb when one does not occur in the input string may prove incorrect in sentences with heavy-NP shift. (example taken from...
of a right-adjacent direct object to a transitive verb (*destroy*) signals that the structure built so far must be revised. This is termed relicensing and it cancels the applicability of the OLLC as there is no source or target position – the whole structure must be rebuilt. For the OLLC to be valid, immediate reattachment is required, as in (16).

Note that the verb must be unambiguously transitive in order to circumvent an otherwise expected garden path effect. The verb must require two obligatory arguments, overt or implicit:

(18) a. The bird bought in the store flew away.
    b. The thug murdered on Thursday was pushing up daisies.
    c. The children found in the woods were frozen.

And compare:

(19) a. *The bird bought in the store.
    b. *The thug murdered on Thursday.
    c. *The children found in the woods.

If the verb selects only a single argument, i.e. if it is either unaccusative or unergative,

(20) a. The boat floated down the river.
    b. The horse raced past the barn.

the garden path effect will occur in reduced relative constructions (as demonstrated), other ambiguities being equal.

The requirement of Adjacency is an observation on Case assignment of the accusative case in transitive verbs, reflecting grammatical knowledge about the position in which the direct object should appear in relation to the verb. Its realization should be established for each language under examination. In English, for example, a VO language, Adjacency requires that the direct object be right-adjacent to the transitive verb. In Japanese, an OV language, the direct object is left-adjacent.

Pritchett, 1992a, p. 95):

(i) The spaceship destroyed *e* in the battle [the giant Kzinti cruiser which has been pursuing it for weeks].

Sentence (i) will be first interpreted as a simplex clause (transitive interpretation) and reparsed as a relative NP upon encountering *in the battle*. When the true internal argument appears, re-structuring will be required. This does not violate the OLLC. *Spaceship*, as head of the relative NP, will have no source position and will therefore be free to attach as a simple clausal subject just as is any unattached element.
Perhaps it is useful now to summarize the model. Theta Attachment is the principle that guides the parser in its attachment decisions. The OLLC is a structural principle on permissible reanalysis. Descriptively, one of the OLLC’s primary predictions is that reanalysis of an argument as a lower or co-argument may be acceptable whereas reanalysis as a higher argument (raising) is illicit. The OLLC is not applicable to any subconstituents of a reanalyzed category, which are themselves not immediately relicensed, as in sentence (17), or internally restructured, as in the case of ‘permitted reanalysis’ in (15).

Having completed the presentation of the Thematic Transparent Model, I will proceed to recapitulate various objections to the OLLC and Theta Attachment in Section 1.5. I will present these criticisms and show that they do not necessitate any change in the model.

1.5. Criticisms of the Thematic Transparent Model

The thematic parser as presented consists of two parts: a guiding principle for initial attachment (Theta Attachment), and a reanalysis constraint (the OLLC). Over the years since the presentation of the Thematic Transparent Model, both experimental and theoretical objections have been raised to both parts of the model. In §1.5.1, I will discuss problems that have been brought against Theta Attachment as a guiding principle of human parsing. These objections argue that attachments are executed immediately, before the verb is encountered. My main argument is that none of the objections undermine Theta Attachment, and the experimental evidence is compatible with the thematic parser.

In §1.5.2, I will discuss two objections against the OLLC from English and Japanese and suggest solutions to these problems.

In Section §1.5.3, I will discuss experimental data that a transparent approach to language could be at fault and argue that these arguments misinterpret the Thematic Transparent Model.

Finally, in section §1.5.4, I will discuss the argument that frequency accounts are an alternative to principle-based parsing (such as the thematic parser). I will show that frequency based accounts are neither sufficient nor necessary to explain garden path effects.

1.5.1. Problems with Theta Attachment

In strict head-final languages (Japanese, Korean, Assamese, Turkish etc.), the verb appears at the end of a clause. The same is valid to non-strict (mixed) verb-final languages (Dutch, German – these permit verbs to follow nominal arguments, but typically precede clausal arguments). According to the Thematic Transparent Model, the NPs preceding the verb should be
stored until the verb appears. Only then is it possible to make a thematic attachment between a verb and a noun. Such are head-driven parsers. This seems to be generally true, as Japanese speakers, for example, do not encounter particular constructions that cause consistent and severe processing difficulties comparable to English garden path sentences (Mazuka et al., 1989).

Nonetheless, a delay in attachment (until the verb appears) was criticized and very specific evidence was brought for 'immediate attachment', even before the verb, from head-final languages (Aoshima, Phillips and Weinberg, 2004, Bader and Lasser, 1994, Kamide and Mitchell, 1999, Miyamoto, 2002, Tanenhaus, Garnsey and Boland, 1990). Although these data are specific and small in scope, they were taken to be the final blow to the core of head-driven parsers and suggested also that Theta Attachment was incorrect.

Before proceeding to discuss objections to Theta Attachment, I would like to mention that Chapter 2 is wholly dedicated to answer the question whether Theta Attachment indeed guides attachments. The chapter provides evidence in favor of Theta Attachment (in Dutch) as a guiding principle of human parsing. The results reported there are compatible only with Theta Attachment and the thematic parser, which assumes it.

The remainder of §1.5.1 is structured in the following way. In §1.5.1.1, I will present experimental evidence from head-final languages against Theta Attachment. In §1.5.1.2, I will put forward evidence in favor of Theta Attachment, based on head-final languages, as well. In §1.5.1.3, I will provide an account of the data discussed in §1.5.1.1. In §1.5.1.4, I will discuss another type of argument against Theta Attachment: that the parser does not consider thematic information in its initial attachment decisions. Further, I will show how thematic information is used by the thematic parser. In §1.5.1.5, I will provide experimental evidence for the use of thematic information during parsing.

1.5.1.1. Head-final languages: ‘immediate attachment’

An example against delay in attachment until the verb and in favor of ‘immediate attachment’ is Bader and Lasser’s (1994) experiment in German. Consider the following sentence:
(21) dass sie (Nom/Acc) nach dem Ergebnis zu fragen
    that she/her for the result to ask
    tatsächlich erlaubt hat / worden ist
    indeed permitted has / has been
    ‘that she indeed has permitted to ask for the result’ (has)
    ‘that to ask her for the result has indeed been permitted’ (has been)

The sentence contains a feminine singular pronoun *sie* locally ambiguous between nominative and accusative case. *Sie* can be either attached as the object of the infinitival verb ‘fragen’ *to ask* (accusative reading), or as the subject of the finite verb ‘erlaubt’ *permit* (nominative reading). The auxiliary at the end of the sentence resolves the ambiguity. If ‘hat’ *has* appears, the nominative analysis turns out to be correct, while if ‘worden ist’ *has been* turns up, the correct reading is the accusative one.

Bader and Lasser predicted that if the parser waits to the verb in its attachment decisions, then *sie* will be initially attached to ‘fragen’ *to ask*, the first occurring verb, in immediate satisfaction of Theta Attachment. If the parser makes ‘immediate attachments’, before verbs appear, it will immediately assume a subject position after the complementizer *dass*, without waiting for a licensing verb.

Using this manipulation in a self-paced reading task, Bader and Lasser found a slowdown at the final auxiliary (*ist*). This suggests that *sie* is assumed to be the subject of the final verb but not of the first. They mention that this finding is incompatible with Theta Attachment.

Kamide and Mitchell (1999) report a preference for ‘immediate attachment’ in Japanese, too, prior to the appearance of a verb:

(22) Kyooju-ga gakusee-ni toshokansisho-ga kasita mezurasii
    Professor-Nom student-Dat librarian-Nom lent unusual
    komonjo-o miseta
    ancient manuscript-Acc showed
    ‘The professor showed the student the unusual ancient manuscript which the librarian had lent.’ (High Attachment)
    Or
    ‘The professor showed the unusual ancient manuscript which the librarian had lent the student.’ (Low Attachment)

The dative-marked NP ‘gakusee-ni’ *student-Dat can be attached either to the subordinate verb ‘kasita’ *lent* or to the matrix verb ‘miseta’ *showed*. The first option was termed ‘low attachment’ and the second option ‘high attachment’. The ambiguity in the sentence is global (it is not a garden path)
in the sense that the grammar does not force either attachment to be confirmed at any point in the sentence.

They compared sentences such as (22) with unambiguous versions in a self-paced reading task. In one unambiguous condition, the dative NP could only be attached to the first verb ‘kasita’ *lent*. A slowdown was observed at the second verb. In the second unambiguous condition, the dative NP could only attach to the second verb ‘miseta’ *showed*, but no slowdown was detected at any point. The outcome is incompatible with Theta Attachment, because it suggests that readers do not attach the dative NP when the first verb is processed, in compliance with what is claimed to happen in German.

In addition, Miyamoto (2002) considers evidence from a self-paced reading task that there is no delay in attachments.

(23) Ofisu-de shokuin-ga kakarichoo-o ocha-o
   Office-Loc employee-Nom manager-Acc tea-Acc
   dashita Josei-ni teineini shoukaishita
   served woman-Dat politely introduced

   ‘At the office, the employee politely introduced the manager to the woman who served the tea.’

The example involves a sequence of two accusative NPs, each belonging to a different clause – a situation that can arise in Japanese. A significant slowdown in residual reading times was found at *och-o* (tea-Acc), the second accusative NP, compared to controls (see example (24)).

(24) Ofisu-de shokuin-ga kakarichoo-ni ocha-o
   Office-Loc employee-Nom manager-Dat tea-Acc
   dashita Josei-o teineini shoukaishita
   served woman-Acc politely introduced

   ‘At the office, the employee politely introduced the woman who served the tea to the manager.’

Miyamoto takes the slowdown to indicate that a clause boundary is built-up, even before the initial verb is encountered. So the slowdown is caused by the need to construct this new clause. Miyamoto concluded that the “… results support the proposal that NPs in verb final clauses are associated incrementally within a partial interpretation even when no verb has been processed” *(ibid., p. 339)*. He adds that “the head-driven explanation is challenged” in that “there is intuition that the NPs … are partially interpreted before the verb is read” *(ibid., p. 312)*, and that “attachment can proceed even if the exact lexical information about the verb is not known” *(ibid., p. 313).*
I will discuss a solution to these specific problems in §1.5.1.3. In the following subsection §1.5.1.2, I present evidence from Dutch, which is similar in structure to German, in favor of Theta Attachment, and against ‘immediate attachment’. I will supplement the Dutch data with experimental evidence from Japanese for delay in attachment, at least until the verb.

1.5.1.2. Evidence in favor of delay

Let us note first that the hypothesis of ‘immediate attachment’ would run into problems in other areas of head-final languages. The evidence comes from relative clauses in Dutch. These are always ambiguous between subject reading and an object reading for the relative pronoun, due to the verb-final character of the language in embedded clauses. Their status is potentially ambiguous until the disambiguating verb is encountered. Consider examples (25) from Mak (2001), who conducted a self-paced reading and an eye-tracking experiment:

(25)

a. Vanwege het onderzoek moeten de inbrekers, die de bewoner beroofd hebben, nog een tijdje op het bureau blijven.
   Because the investigation should the burglars who the occupant robbed have still some time at the police office stay
   ‘Because of the investigation, the burglars who robbed the occupant will have to remain at the police office for some time.’

b. Vanwege het onderzoek moet de bewoner, die de inbrekers beroofd hebben, nog een tijdje op het bureau blijven.
   Because the investigation should the occupant who the burglars robbed have still some time at the police office stay
   ‘Because of the investigation, the occupant who the burglars robbed will have to remain at the police office for some time.’

In (25), an ‘immediate attachment’ approach would predict a preference to assume a gap (the trace of the relative operator) in the subject position of the relative clause immediately following the relative pronoun (Active Filler Strategy, as in Frazier (1987b)). The type of relative clause (object or subject relative clause) is disambiguated by the auxiliary. Since ‘immediate attachment’ predicts that the initial analysis will always be a subject interpretation for the relative pronoun, the sentence where the auxiliary

---

15 I am grateful to Iris Mulders for pointing out these examples to me.
matches the relative pronoun in number is predicted to be read faster there. Mak’s results are compatible with such a prediction.

However, Mak shows that when the animacy properties of the relative pronoun and the NP following it are manipulated, the general preference for a subject relative clause disappears. More importantly, when the NP following the relative pronoun is replaced by a pronoun, it turns out that the object relative clause interpretation is preferred. This suggests that, in fact, the parser must delay in its thematic attachment decision at least until the verb (Mak, 2001). Consider:

(26)

a. Ongerust kijkt de hardloper, die jullie in het park heb je, naar de donkere lucht.
   Worried looks the jogger who you (Pl) in the park has at the dark sky
   ‘Worriedly, the jogger, who greeted you in the park, looks at the dark sky.’

b. Ongerust kijkt de hardloper, die jullie in het park heb je, naar de donkere lucht.
   Worried looks the jogger who you (Pl) in the park have at the dark sky
   ‘Worriedly, the jogger, who you greeted in the park, looks at the dark sky.’

(26) is again ambiguous until the auxiliary, and the pronoun ‘jullie’ you (Pl.) is ambiguous between nominative and accusative interpretations. The structure is disambiguated towards the object relative clause at the auxiliary, because reading times were shorter when the pronoun was disambiguated towards the nominative interpretation.\textsuperscript{16} (26) shows that, in Dutch, the parser does not assume a subject position before it is licensed, as suggested by Bader and Lasser in §1.5.1.1. This goes against ‘immediate attachment’ models, since the parser has to wait \textit{at least} until after the NP following the relative pronoun to disambiguate the structure, correctly predicting no difficulty in both cases.

\textsuperscript{16} Kaan (1997) tested embedded \textit{wh}-questions in Dutch (self-paced reading tasks, mostly) and found a general preference for the \textit{wh}-element to be interpreted as the subject. However, when the \textit{wh}-element was followed by a pronoun, the \textit{wh}-element was preferably attached as the object. This is compatible with Mak (2001) discussed in the main text.
There is evidence from Japanese that thematic attachments do occur only at the verb position. This type of evidence also goes against accounts that assume 'immediate attachment'.

Aoshima, Phillips and Weinberg (2004) show in a self-paced reading task that Japanese readers associate a fronted wh-phrase (which student) with the first verb they encounter, rather than with the structurally highest verb:

(27) a. Question particle

\[
\text{Tannin-wa koocyoo-ga dono-seito-ni hon-o class teacher-Top principal-Nom which student-Dat book-Acc yonda-ka tosyositu-de sisyo-ni iimasite. read-Q library-at librarian-Dat said}
\]

‘The class teacher said to the librarian at the library which student the principal read a book for.’

b. Declarative complementizer

\[
\text{Tannin-wa koocyoo-ga dono-seito-ni hon-o class teacher-Top principal-Nom which student-Dat book-Acc yonda-to tosyositu-de sisyo-ni iimasite-ka? read-Decl library-at librarian-Dat said-Q}
\]

‘Which student did the class teacher say to the librarian at the library that the principal read a book for?’

Sentence (27a) was compared with a similar sentence (27b), where the question particle (\text{ka}) was replaced by a declarative complementizer (\text{to}), and the question particle was suffixed to the final verb. Reading times were significantly higher at the first verb \text{read} when there was a declarative complementizer. This reflects the preferences of Japanese speakers to attach the fronted wh-phrase with the first verb. It is for this reason that a slowdown was measured when a declarative particle was found: Japanese speakers had to revise their attachment of the wh-phrase to the first verb. Because wh-phrases were marked dative, and the verbs were mono-transitive, the effect is not due to the theta grid of the first verb, but rather reflects the requirement of the wh-phrase to receive thematic interpretation as soon as possible. This finding is therefore consistent with immediate satisfaction of Theta Attachment.

Relatedly, there is evidence from English against the strong hypothesis that attachments are immediately head triggered. Consider the following example (28):
(28) The patient persuaded the [NP doctor that disliked him], PRO, to leave.

(29) ¿The patient persuaded the [NP doctor that he agreed with], PRO, to leave.

When *persuaded* is encountered, the parser retrieves its maximal theta-grid, having three thematic roles <AGENT>, <THEME>, <PROPOSITION>. In a strict head-driven approach, a CP would be created already when the parser encounters the complementizer *that*, being a head. However, the lack of an overt subject in the substring *that disliked him* in (28) bleeds the complement clause interpretation, leaving only the subject relative clause reading (note the contrast with (29), where the substring *that he agreed with* lacks such information as in (28) leading to a garden path effect). The unproblematic status of sentence (28) shows that attachment does not appear to occur immediately given the head of CP, *that*. Theta Attachment differs from such strict head-driven approaches in that only a lexical-thematic head triggers attachment. Clauses are not actually licensed until the occurrence of a verb, their ‘true’ head, which explains why this sentence does not pose processing difficulty.

1.5.1.3. Accounting for the putative problems for Theta Attachment

Returning to the German example (21), note that ‘immediate attachment’ predicts any element that occurs after the complementizer *dass* to be attached as subject. This does not appear to be correct. When *sie* was replaced by the unambiguous *ihn* (accusative, male, singular), it was not available for attachment to the final verb. In an approach that supports ‘immediate attachment’, *ihn* should be unavailable for attachment with the embedded infinitive, and clauses with *ihn* should lead to parsing problems when the auxiliary is *ist*. However, no processing difficulty was found in such cases, as Bader and Lasser point out. ‘Immediate attachment’ seems to make an incorrect prediction with regard to the attachment of *ihn*.

Bader and Lasser account for this discrepancy by stating that if *ihn* were to attach to the final verb, the structure would be highly marked, and therefore it is not preferred:

\[
(30) \text{[dass Object [Infinitival clause] Subject Verb]}
\]

However, it is not clear from Bader and Lasser’s discussion how this would work exactly: why is the accusative attachment of *sie* not rejected on the same ground as that of *ihn*, given that *sie* in its accusative interpretation also correlates with the structure in (30).

The conclusion is that the argument of ‘immediate attachment’ is not fully supported by this type of experimental data. What do the data show? They show a lack of processing difficulty when *ihn* is found, and processing
difficulty when *sie* cannot be attached as subject of the final verb. These findings are accountable by Theta Attachment as follows.

Recall that the parser assumed here is transparent, and must therefore satisfy all the principles of syntax maximally at every point during parsing (Pritchett’s quote on page 8). This is Pritchett’s Generalized Theta Attachment (Pritchett, 1992a, p. 138). One of the principles of syntax is the EPP (the Extended Projection Principle), which means that a clause must project a subject position.

In the German example (21), ‘*sie*’ *she* is (potentially incorrectly) identified as nominative, but may not be attached to ‘zu fragen’ *to ask* since an infinitival cannot check nominative case. The first opportunity to satisfy the EPP appears at the final verb in the active form. It is possible at this point to satisfy this grammatical principle, and decrease the pressure on the parser of satisfying syntax. This attachment turns out to be incorrect then, when the passive auxiliary is encountered, and is reflected by a slowdown.

It is absolutely crucial to recognize that syntactic principles are increasingly satisfied to the degree that each element in the local representation is licensed, and, beyond this, the greater the number of licensing features (e.g. EPP, theta-roles) can be discharged. This is not equivalent to the claim that the parser attempts to construe each local string as globally grammatical, though at times it may have the same effect.

It should be emphasized that the satisfaction of the EPP does not depend on the existence of auxiliaries in the input. In strict head-driven models (which are not thematically-driven), a subject position would be projected when an auxiliary is met to immediately satisfy the EPP. However, in sentence (31),

(31) Below the stairs would it be possible to store the bike?

processing difficulty would be erroneously predicted, because, when *would* appears, *the stairs* should be readily attached to a subject position projected by *would*: \[tp \text{Below, } [TP \text{ [np the stairs] would}] \].

What happens in (31) requires properly understanding the thematic model and the OLLC, which characterizes analyses that the parser cannot perform, rather than operations that actively cause it to fail. Under this view, the required reanalysis (where *the stairs* should move from the direct object position of the preposition to the subject position of *would*) can never occur, since *would* has no thematic content. Rather, the parser continues its automatic operation, and, when a pronoun appears, it is capable of serving as the subject of *would* in an inversion construction, salvaging the parse. In sentence (32) in contrast, the situation continues to deteriorate with no subject appearing and eventually a costly reanalysis is required.
(32) ¿Below the stairs would collapse.

The Japanese sentence (22) is easily accounted for by the same EPP requirement that accounts for the German sentence (21). There is overwhelming syntactic evidence that an NP-ni NP-ga sequence, found in (22), is locally ambiguous between a dative-nominative and a double subject analysis (examples and discussion follow Pritchett, 1992a, section 5.2.4, page 151):

(33)  a. Dative -ni
  Rex-ni John-ga hanasita.
  Rex-Dat John-Nom spoke
  ‘John spoke to Rex.’

  b. Subject -ni
  John-ni nihongo-ga wakaru.
  John-Dat Japanese-Nom understand
  ‘John understands Japanese.’

The parser attempts to satisfy the EPP when the first verb lent is encountered in (22). However, given the structural ambiguity of the -ni particle, and the sequence itself (containing actually three possible candidates to occupy the subject position: the two -ga marked NPs and the -ni marked NP), the parser may delay its syntactic decision to the imminent second verb. This accounts for the alleged ‘attachment’ of the – ni element to the final verb.

With regard to Miyamoto’s example (23), let us notice that the sole evidence for the putative construction of a new clause is the finding of slowdown in this area. However, a slowdown may have many reasons and it is not clear that the slowdown found at the second accusative NP means there is pre-verbal attachment, or that a clause is being built, as Miyamoto claims. The slowdown could be related to a failure to meet an expectation: after an accusative marked NP, the parser expects to find a verb. Recall that Japanese respects Adjacency, although less directly because of word order flexibility. In informal conversations, the accusative marker (normally suffixed to direct objects) can be omitted from direct objects, but grammaticality is degraded if the direct object without case marker is not immediately followed by the verb (Saito, 1985). When the accusative is not adjacent to the verb, there is a measurable effect in Japanese (Ueno and Kluender, 2003, among others). The relevance of Miyamoto’s experiment to Theta Attachment is therefore moot.
1.5.1.4. Thematic information is not used by the parser in initial resolution of attachments

One of the longest debates in psycholinguistic literature revolves around the question whether thematic information, mostly relating to animacy, may override purely structural considerations.

Let us trace back the steps of this debate, as it is relevant for establishing the reliance of the thematic parser on thematic information.

Ferreira and Clifton (1986) showed disruption of reading in the region by the lawyer in both (34a) and (35a), compared to their respective unambiguous counterparts (34b) and (35b) in an eye-movement experiment.

(34)  

a. The defendant examined by the lawyer turned out to be unreliable.

b. The defendant that was examined by the lawyer turned out to be unreliable.

(35)  

a. The evidence examined by the lawyer turned out to be unreliable.

b. The evidence that was examined by the lawyer turned out to be unreliable.

According to models that rely on structure, but not on thematic information for the initial resolution of attachments, sentences (34a) and (35a) should equally give rise to a garden path effect. In (34a), the defendant is given a main clause analysis in which it is the subject and <AGENT> of examined. The phrase by the lawyer and the following main verb phrase turn out to be inconsistent with the main clause analysis, forcing reanalysis. This slows down reading (related to a garden path effect) compared to (34b), in which the phrase that was examined is unambiguously a relative clause that modifies the first noun, defendant.

Ferreira and Clifton (1986) reasoned that if readers immediately used meaning and plausibility, not only structural knowledge, to guide their initial interpretation of a sentence, then the disruption of reading would disappear in (35a), with an inanimate initial noun. Evidence is not a plausible subject of the verb examined, and this should block analysis in which it is the subject. Rather, since it is a plausible <THEME> of examined, it will be initially taken as the head of a (reduced) relative clause. No reanalysis will be needed when the eyes reach the by the... phrase and no disruption will take place. Ferreira and Clifton (1986) observed very comparable amounts of disruption in (34a) and (35a). They argued that this meant that the evidence was initially taken to be the subject of the verb examined, contradicting an apparent prediction of meaning/plausibility based models and supporting a model in which grammatical constraints are honored first even when they clash with plausibility.
Trueswell et al. (1994) (others also experimented on this matter, e.g. Stowe (1989)) challenged the above-mentioned findings and claimed that plausibility did affect attachment decisions. However, Clifton et al. (2003) (and others, e.g. Lipka (2002)) reported again disruptions in the reduced relatives whether the sentence’s subject was animate or inanimate (i.e. irrespective of animacy). They noted that “while it is certainly possible that syntactic preferences can be eliminated in some domains, it appears that reduced relative clause domain… is not one of them” (p. 331).

Before addressing these findings, let me first discuss how the thematic parser uses thematic information during on-line parsing. I will then continue to state the predictions of the thematic parser with regard to sentences (34a) and (35a), and suggest that the predictions are compatible with Ferreira and Clifton’s findings.

The satisfaction of the Theta Principle involves a head that selects an argument under feature identity (see page 13). The parser must take into consideration the ‘fit’ between the label of the identified argument and the label of the role selected by the head. Furthermore, it must determine the fit between semantic features of the thematic role and the argument. Relevant questions are how this information is encoded and what part of it is available for the syntactic component of the parser.

Thematic labels are outside of grammar (Chomsky, 2004), but in order to determine what kind of arguments a head can be associated with, the parser must be able to access minimal information about conceptual structure. This means that in order for a verb, corresponding to a verbal concept, to be associated with its selected arguments, there must be access to conceptual structure. An example of such encoding is provided within the Theta System (Reinhart, 2000, 2001, 2002, Reinhart and Siloni, 2005b). Within this system, thematic labels can be encoded by two binary features \([\pm c, \pm m]\): \([c]\) stands for *cause change*, \([m]\) stands for *mental state*, \([+]\) denotes the relevance of such feature, \([-]\) its irrelevance. To give an idea how these features are used in encoding, the thematic labels of the different feature clusters are given below:
As a concrete example, the verb *drink* selects an *AGENT* encoded as [+c, +m]. Under feature identity, this role may be realized by *John* (as in *John drinks*) but not by *the table* [-c, -m] (*The table drinks*). This is so, because only an animate NP, like *John*, can satisfy the [+m] feature included in the cluster that the verb selects. We should note, however, that nothing in the present discussion hinges on this specific assumption regarding how 0-features are assigned and checked, and the same can be captured under other implementations of theta assignment.

In the study of parsing performance, and especially in the case of a thematic parser, semantic restrictions play a part during parsing. Because the parser is guided by Theta Attachment and the Theta Principle, it must incorporate such semantic restrictions. In examples (34a) and (35a), semantic restrictions may guide the parser in making the correct attachment decision initially. In (35a), the verb *examine* requires [+c, +m], which cannot be awarded to *evidence*, because it is inherently inanimate. There will be no attachment and the correct grammatical representation will be obtained. In (34a), *the defendant* may assume the [+c, +m] role of the verb *examine*. Note, however, that sentence (34a) falls under the type discussed in section §1.4.3.2. This is an instance of an obligatory transitive verb, occurring without an immediate accusative complement. Upon encountering the phrase *by the lawyer*, the parser realizes that the direct object of *examine* (being obligatory transitive) will not appear, because of the Adjacency requirement operative in English. As we saw there, in such cases, relicensing is required which cancels any source or target position and is therefore not costly. Hence, (34a) contains uncostly relicensing, and (35a) is supposed to be easy. Thus although there is a difference in the derivational steps in these sentences, it may not lead to a visible difference in processing ease. These predictions are compatible with Ferreira and Clifton’s (1986) and Clifton et al. (2003) results because although there are disturbances in the parsing of both sentences (34a) and (35a), they are not predicted to be significant – the OLLC is not violated and there is no severe garden path. As I mentioned, the experimental results themselves are debated. Possibly, the source of...
disagreement in the interpretation of the findings stems partially from the fact that relicensing may still take slightly longer than a case that does not require relicensing.

There is further evidence that θ-features do play a role in making attachment decisions. In (37), the hammer will not be initially attached to drank. This is so because of finer selectional restrictions than the ones we have examined so far. Although the hammer can serve as a \(<\text{THEME}>[-c,-m]\), the verb drink specifically requires an NP denoting liquid to fulfil this role.

(37) After Mary drank the hammer fell off the stool.

Had the hammer been attached to drank, a garden path effect would have been predicted, contradicting actual performance. In this respect, the thematic parser is consistent with parallel approaches and the relevant psycholinguistic evidence that people are sensitive to selectional restrictions (cf. Tanenhaus, Carlson and Trueswell, 1989). In the next subsection, I turn to more such evidence.

1.5.1.5. Evidence in favor of use of thematic information

This subsection contains a non-exhaustive overview of findings from psycholinguistic experiments in various modalities that thematic information (mainly animacy) is being used in on-line processing. The overview further establishes the status of Theta Attachment and the Theta Principle as principles of the parser.

In Event-Related Potential (ERP) experiments, where changes in brain electric activity are monitored, it was found that semantic violations modulate the amplitude of a distinct wave form 400 msec after detection of a semantic mismatch (Osterhout, 1994, Osterhout and Holcomb, 1992, Osterhout, McLaughlin and Bersick, 1997, Osterhout and Mobely, 1995). For instance:

(38) The boat sailed down the river and barked during the storm.

The selectional restriction in (38) elicited an N400 (N = Negative) wave; showing that semantic restrictions, defined as ‘thematic fit’ or thematic ‘Match’, are being used during sentence processing.

Agrammatic patients were also reported to be sensitive to ‘thematic fit’. Error rates were higher in agrammatic patients on sentences where semantic constraints conflicted with the syntactically based assignments, even in sentences with canonical (SVO) word order (e.g., #The deer shot the hunter) (Safran, Schwartz and Linebarger, 1998).

Further studies of agrammatic patients have shown that patients can distinguish between verbs with different thematic structure. Their error
patterns could be explained by a theory that assumed a lexicon that encodes thematic roles (Grodzinsky, 1995a, b). These findings are consistent with an approach that theta-roles are encoded in the lexicon, unlike adjuncts (Boland, 2005a). Boland (2005b) also demonstrated in an eye-movement experiment (visual world paradigm) that argument structure focuses listeners’ attention on relevant aspects of a visual scene. For example, when listeners heard a dative verb, they made more anticipatory fixations on photographs depicting potential recipients.

In a cross modal sentence priming study, it was demonstrated that verbs primed for core thematic roles such as <AGENT> and <THEME> (Ferretti, McRae and Hatherell, 2001); for example arresting primed cop, and criminal, but not locations (swam – ocean). The study therefore implies that ‘thematic fit’ affects priming: having retrieved the theta grid of arresting and the features it selects, an upcoming argument will match with the retrieved features. There is no problem of ‘look-ahead’ as the expectation is based on given data, i.e. the retrieved θ-features.

There can at least be no doubt that lexical information such as the verb’s thematic grid is being considered during sentence parsing, and that the parser employs the concept of thematic fit. A thematic approach does not ignore these factors, but asks once a certain thematic grid has been retrieved, what is its interaction with knowledge of grammar.

Given the overwhelming experimental support to the use of theta-roles, it should be additionally noted that most principle-based models assume a thematic module, where thematic information is being handled (Crocker, 1996, Fodor and Inoue, 1998, Frazier, 1987a, Gorrell, 1995b). It is only that Pritchett (1988, 1991, 1992a, 1992b) has adequately developed it, giving it central stage.

1.5.2. Problems with the OLLC

In this section, I will discuss problems that were raised against the OLLC and the solutions suggested for these problems. In §1.5.2.1, I will provide an example raised against the OLLC from Japanese, a head-final language. I will proceed to present a solution within a head-driven thematic approach. §1.5.2.1 is not meant as an exhaustive discussion of the full range of problems with the OLLC in head-final languages. A comprehensive discussion and solution of the problems raised for Japanese in particular and head-final languages in general are found in Mulders (2002). In §1.5.2.2, I will discuss an argument that seems to show that raising (a violation of the OLLC) can be easy in English. I will argue that there is no direct evidence that this is so.

The discussion in Section §1.5.2 aims to emphasize the point that the OLLC, just like Theta Attachment, is valid, even for head-final languages.
1.5.2.1. The OLLC is not valid for Japanese

Mazuka and Itoh (1995) argue that in Japanese immediate reanalysis violating the OLLC is carried out, but is not costly.\(^{17}\) Consider their example:

\[(39)\] Nakamura-ga tyuuko-no pasokon-o katta
Nakamura-Nom second hand computer-Acc bought
toki syuuri-site-kureta.
when repaired (for me).

‘When I bought a second-hand computer, Nakamura repaired it for me.’

When the final matrix verb is encountered, there should be immediate reanalysis of Nakamura-ga from the subject position of the initially built CP ‘Nakamura-ga tyuuko-no pasokon katta toki’ when Nakamura bought a second-hand computer to the subject position of the matrix verb repair for me. This ‘raising’ obviously violates the OLLC, but the sentence does not invoke processing difficulty – the OLLC allegedly makes the wrong prediction with regard to this sentence.

‘Raising of Nakamura-ga is necessary because, as Mazuka and Itoh point out, the verbal construction ‘syuuri-site-kureta’ repaired for me has the idiosyncratic property that it cannot take pro as its subject (which is interpreted as the speaker).\(^{18}\) The subject of the embedded clause must be the subject of the matrix clause, and pro must Merge in the embedded clause’s subject position.\(^{19}\)

Mulders (2002, 2003) discusses this and various other Japanese counterexamples to the OLLC, and suggests an alteration to Pritchett’s original OLLC to accommodate Japanese data. She suggests that raising is possible when the source position is at the leftmost Specifier (‘left edge’) containing overt material, and when the target position requires the constituent to leave its own phase (phases are syntactic constructs. For the purposes at hand, a phase is either a CP or a VP). The NP Nakamura-ga in (39) is reanalyzed from the leftmost Specifier position of the temporal

\(^{17}\) Hirose (1999) and Miyamoto (2002) argue that Mazuka and Itoh’s paper is partial and does not provide the full picture of the data. I will assume, though, that Mazuka and Itoh are correct, at least about reports of processing difficulty or lack of them, following Mulders’ (2002) policy on this issue.

\(^{18}\) One should note that the determination of the speaker of the sentence plays an important part in circumventing processing difficulties. An elaborate discussion is found in subsection §3.10.1.

\(^{19}\) It is possible to interpret the sentence with a pro in the matrix subject position referring to a third person, but this reading is hardly obtained without constraining context.
adjunct [Spec, TP], and must abandon its position within the CP phase, as demonstrated in Tree 10.

Mulders’ revised OLLC therefore correctly accounts for the lack of a garden path effect in this case.

The main point to be gleaned out of this concise discussion is that Japanese data are accommodated within the thematic parser, as demonstrated in Mulders (2002), contra Mazuka and Itoh’s objections. Although Mulders suggested revising the OLLC relying on specific Japanese cases, the original OLLC suggested by Pritchett still makes the correct predictions over a vast range of cross-linguistic data and I will refer to it in the remainder of this work for ease of exposition.

1.5.2.2. The OLLC is not valid for English

Fodor and Inoue (1998) argue that in English, an OLLC violation is available without processing cost. They present a sentence that “…disqualifies Pritchett’s OLLC” (p. 118):

(40) While the boy scratched the little cat and the big hairy dog yawned loudly.

Sentence (40) contains the same type of object-subject ambiguity as example (14). Fodor and Inoue observe that initially, the entire coordinate NP is construed as the object of the first verb.
(41) While the boy scratched [the little cat and the big hairy dog] [yawned…]

However, only the second conjunct of the coordinate structure is raised. The resulting string is ungrammatical, or, at least incomplete, since the whole string has been structured as a while-clause with internal coordination, and the main clause is missing:

(42) While [[the boy scratched the little cat] and [the big hairy dog yawned]]…

The string would be acceptable only if the input contained another clause:

(43) While [[the boy scratched the little cat] and [the big hairy dog yawned loudly]] Kim slept.

The latter sentence is not difficult to understand despite its complexity. Fodor and Inoue argue that the parser remains in representation (42), “without reaching” (p. 118) the grammatical interpretation:

(44) While [the boy scratched] [[the little cat and the big hairy dog yawned loudly]].

Fodor and Inoue’s main point is that “raising a constituent can be easy” (p. 118), as appears to be evident by the lack of difficulty of raising the DP the big hairy dog from its source position within the direct object position of the verb scratch to the subject position of yawned (easy obtainment of (42)). They argue that this is a problem for the OLLC, since it would erroneously predict difficulty in this particular instance.

Is it so that the (alleged) violation of the OLLC in (42) is not costly? When while is removed from (40), there is a garden path, though a mild one (Meseguer, Carreras and Clifton, 2002):

(45) The boy scratched the little cat and the big hairy dog yawned loudly.

Example (45) contains the same operation as in (42) (where a coordinated NP must be broken and then a part of it raised), but, given Fodor and Inoue’s analysis, should be entirely unproblematic – which is not the case. Therefore, there is no direct evidence that there is no difference in the cost of reanalysis between (42) and (45).

I would like to stress that the OLLC predicts conscious reanalysis. The parser is unable to perform the reanalysis necessary to obtain a grammatical representation and requires the invocation of a higher cognitive mechanism. Garden paths represent unrecoverable parsing errors which cannot be corrected without the conscious invocation of higher rational and nonautomatic cognitive processes. The OLLC does not predict the perceived degree of reanalysis. For example, it does not explain why (45) poses only a
mild garden path. This point relates to one of the criticisms of the thematic parser and the OLLC: that it ignored factors related to actual performance that may ameliorate or exacerbate the experience of the garden path effect. It was often wrongly conjectured that if the model predicted a violation of the OLLC, then individuals would (consciously) experience an identical severe degree of difficulty across the board.

However, it should be realized that the conscious experience of the garden path effect may vary in its experienced severity. There is a distinction between a conscious experience and the degree with which it is experienced. Thus, the OLLC predicts a conscious experience without necessarily committing to the degree in which it is experienced or whether it will be perceived.

1.5.3. Strong transparent approaches to parsing do not correspond with actual performance

Very convincing evidence has been brought to show that readers reach only ‘good enough’ interpretations (i.e. ungrammatical) of garden path constructions (Ferreira, Bailey and Ferraro, 2002, Ferreira, Christianson and Hollingworth, 2001, Ferreira and Stacey, 2000). It was shown, for instance, that participants correctly believed after reading the sentence ¿While Anna dressed the baby spit up on the bed that the baby spit up on the bed, and incorrectly yet confidently believed that Anna dressed the baby. These results were argued to imply that ‘garden-path reanalysis’ was not an obligatory process and that thematic roles initially assigned for the subordinate clause verb were not consistently revised. Essentially, subjects do not have to reach a grammatical representation. This implies that the parser does not have to be strongly transparent.

These findings are explained within the thematic parser. The OLLC is a constraint on possible reanalysis. As it is stated, it defines certain structural reanalyses to be impossible for the parser, and as such does not contain the obligatory execution of reanalysis. Reanalysis may not take place at all, and a grammatical representation may not be reached. It is only that in cases of garden path constructions, if structural reanalysis of the sort constrained by the OLLC is carried out, the outcome is costly.

Note that comprehension is equated with the obtainment of a grammatical representation; however, this is somewhat misleading. People may comprehend a sentence that is not grammatical (Me likes chocolate, for instance), or not comprehend a sentence that is grammatical (as in the case of center-embeddings). Maintaining the idea that a grammatical representation does not have to be obtained would result in the loss of an important theoretical distinction for those who do experience a problem in comprehension.
A strong transparent approach, which heavily relies on grammar, may be falsified by showing that there is an error in the principle that the parser does not relate lexical entries with grammar – a highly unlikely crocket.

1.5.4. Frequency is an alternative account for the thematic parser

Frequency-based or constraint-based accounts argue that comprehension is achieved through parallel satisfaction of multiple probabilistic constraints, including constraints from lexical representations. The mental lexicon accommodates extensive grammatical information, including tense, number, person, gender, and thematic structure among others. These models suggest that the parser may be affected by non-structural factors such as the verb’s meaning frequency, which may affect the evaluation of the alternative structural representation of ambiguous sentences (Ni, Crain and Shankweiler, 1996, Ni et al., 1998).

Such accounts are suggested as alternatives accounts of processing difficulty related to garden path constructions. They claim that the data can be fully accounted for by lexical information (Mitchell, 1994), and that lexical frequency has a significant role in syntactic ambiguity resolution (Boland et al., 1995, Ferretti, McRae and Hatherell, 2001, MacDonald, 1993, 1997, MacDonald, Pearlmuter and Seidenberg, 1994a, b, McRae, Ferretti and Amyote, 1997, Spivey-Knowlton, Trueswell and Tanenhaus, 1993, Trueswell, Tanenhaus and Garnsey, 1994, Trueswell and Tanenhaus, 1994).

Giora (2003) explains that difficulty of ambiguity resolution varies with the degree of salience of alternative interpretations available, in compliance with MacDonald, Pearlmuter, and Seidenberg (1994b). The canonical garden path *The horse raced past the barn fell* is more difficult to process than a structurally similar sentence such as *The former mental patients heard here sound unusually sane*, because of the varying degrees of frequency of rival interpretations. The more frequent or salient interpretation of *hear* is transitive (here in its passive form) and it has a weaker competitor in the shape of the less salient intransitive interpretation. *Race* assumes its intransitive meaning more frequently. Thus, there is strong interference of the more salient interpretation (the intransitive reading) with the less salient transitive interpretation. These models suggest that Theta Attachment may not be the main guiding principle of attachment. They also suggest that models that rely on grammar (strong transparency) could be completely replaced by models that rely on probability/frequency.

Are constraint-based models indeed an alternative account to thematic transparent parsing? Ford, Bresnan, and Kaplan (1982) argued that lexical recategorization of the verb takes place on the basis of frequency information and that this is the source of difficulty. Pritchett replied:
“Such a move would be virtually identical to extending the author’s general account of structural attachment heuristics to GP effects.” (Pritchett, 1992a, p. 55)

However, under strict scrutiny it appears that there are some possible differences between a thematic parser and constraint-based models. These data suggest that constraint-based models cannot replace the Thematic Transparent Model.

Pritchett reviewed constraint-based approaches and argued that they failed in their account of garden path phenomena (see §2 in Pritchett, 1992). Let me summarize his major points.

Lexical ambiguity or competition between frequently occurring (salient) interpretations/theta grids is not a sufficient condition for processing failure. For instance, the following sentence involves lexical ambiguity but does not result in processing difficulty or attachment ambiguity:

(46) Ron charged the bank.

The constituents bank and charge are lexically ambiguous, but this fact does not lead to conscious difficulty of the type identified in garden paths. Lexical ambiguity is not a necessary condition for attachment ambiguity, too, or more precisely, syntactic ambiguity:

(47) The chicken is ready to eat.

These sentences have been typically ignored or treated as a distinct subclass, subject (arbitrarily) to independent, albeit generally unspecified, principles.

In the canonical garden path constructions, such as The horse raced past the barn fell, an account based on lexical ambiguity would also predict processing difficulty in a structurally identical sentence:

(48) The tank blown up in the battle was removed from service.

Sentence (48) contains the same type of lexical competition as in the canonical example: blow up assumes an intransitive interpretation, which competes with the transitive interpretation, but this does not lead to processing difficulty. Moreover, there are sentences that do not involve lexical recategorization of the verb, but are still difficult to process:

(49) The cotton fields produce makes warm clothing.

(50) The patient persuaded the doctor that he was having trouble with to leave.

In other words, lexical ambiguity may exist in garden path sentences, but it is neither necessary nor sufficient.
1.6. The garden path effect as local ungrammaticality

As we have seen in the previous section, there is no reason to suspect that the thematic parser is incorrect or has no ‘goodness of fit’. But we may ponder upon the nature of the OLLC. Given that the OLLC predicts illicit raising to be costly, we may inquire as to the motivation of raising. Pritchett does not supply a reason that would motivate raising. He excluded a violation of a grammatical principle to be the motivation, because that would overpredict processing difficulties:

“Theta attachment is a parsing heuristic which resolves local ambiguity by building a structure which maximally satisfies a particular grammatical constraint (or constraints), but this does not in any way imply that there is inherent cost associated with the temporary inability to satisfy the constraint where not grammatically possible. Were such situations hypothesized to be a significant problem for the processor, parsing difficulty would be vastly overpredicted contra the prima facie empirical evidence.” (Pritchett, 1992a, p. 14)

In the following, I will claim that raising is motivated by the local violation of the Theta Principle. The local violation of the Theta Principle serves as an incentive or signal for correction, and raising is the ‘cure’ of the local violation.

The OLLC involves a representation in which an element is removed from a source position and raised to a target position, schematically:
The representation in Tree 11 may be split into two. The splitting of Tree 11 is similar to a computation conducted in derivational analyses. Within theoretical syntax, Reinhart (2006) suggests that under certain circumstances involving violations of some linguistic condition, computing a given derivation requires constructing an alternative derivation in order to determine whether a given step in the current derivation is permitted. This computation is known as reference-set computation.

Similarly, in parsing performance, a local violation of the Theta Principle is diagnosed in the initially obtained representation. This representation is naturally not permitted by grammar. An alternative representation is built, which is permitted by grammar (the ‘repaired’ representation). This sort of local computation may be termed as local reference-set computation.

The representation initially built involves an argument in its source position; the repaired representation has the same argument in the target position:

(51)

<table>
<thead>
<tr>
<th>a. Argument in source position</th>
<th>b. Argument in target position</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Tree diagram a" /></td>
<td><img src="image" alt="Tree diagram b" /></td>
</tr>
</tbody>
</table>

Representation (51a) stands for the structure before raising. Representation (51b) is the repair of (51b), and the grammatical representation.

Consider next a concrete example involving object-subject ambiguity:

(52) ¿After the man left the shop closed.

The argument *the shop* is raised from its object position of *leave* (the source position) to the subject position of *close* (the target position). This raising violates the OLLC and a garden path effect is correctly predicted.

Complying with the split representation provided in (51), we may state the following. The parser initially builds a representation where *the shop is
attached to *leave* in compliance with Theta Attachment. This representation corresponds to (51a). When *closed* is encountered, there is no argument that can be pMerged with the thematically licensed subject position. The grammar informs the parser that a local violation of the Theta Principle has been diagnosed. In other words, a local violation of the Theta Principle is signaled to the parser by grammar. The representation so far obtained is not permitted by grammar. In order to correct this situation and to better satisfy grammar (the Theta Principle), an alternative representation is constructed. The alternative representation repairs the local violation of the Theta Principle, and corresponds to representation (51b).

With regard to the parser, the repaired representation warrants the rejection of its most basic principle of operation: Theta Attachment. The verb *leave* cannot satisfy its maximal theta grid anymore, but the rejection of Theta Attachment is justified on grammatical grounds. The parser is ‘punished’ by grammar for its strict adherence to Theta Attachment: it has led the parser to locally violate the Theta Principle.

Note that not every local violation of the Theta Principle is costly. In the course of parsing, there could be many local violations of the Theta Principle, as Pritchett indicated. A local violation is costly only when the repair operation involves raising – i.e. an operation that violates the OLLC.

It is in this sense that the garden path effect is local ungrammaticality. The garden path effect is the failure to avoid the local violation of the Theta Principle when the cure is raising. The repair or cure operation (raising) is signaled by the local violation of the Theta Principle.

It is important to realize that the ability to omit the direct object (as in the case of *leave*), i.e. the verb’s ambitransitivity, is a crucial criterion for establishing whether raising is to be carried out at all. That is the meaning of the condition *where grammatically possible* for a costly violation of a grammatical principle, as in Pritchett’s quote on page 43.

Since the parser is guided by Theta Attachment, it is forced to initially yield a representation that locally violates the Theta Principle in (52). The cure, which involves raising, becomes available because the rejection of Theta Attachment is grammatically possible. In (52), the verb *leave* can omit its argument *the shop*, as the grammaticality of (53) shows:

(53) The man left.

The omission of an argument that can occupy the <THEME> position is possible because the verb *leave* has a specific lexical property – it is ambitransitive. The <THEME> position is optional: it may be realized or not. It is this lexical property that permits raising: the verb may give up its theta-marked position without loss in grammaticality. Having this
grammatical possibility of ambitransitivity is crucial for raising. Otherwise, this cure would not be permissible by grammar.

We may state now the circumstances under which the local violation of the Theta Principle is costly. Consider Flowchart 1 below.

**Flowchart 1: Deciding whether the local violation of the Theta Principle is costly (the garden path effect)**

The parser identifies a possible local violation of the Theta Principle: there is a theta-role without an argument, or an argument without a theta-role. Grammar, specifically the Theta Principle, insists that the parser repair the situation – the grammar (or the Conceptual-Intentional (C-I) interface)
cannot bear a situation where there is no one-to-one mapping between arguments and theta-roles.

Several cures are possible. One cure was described as ‘permitted reanalysis’ (§1.4.3), when the target position is governed by the source position as in example (15) They gave her photos to Doug.

The cure in (15) was grammatically possible since her was lexically ambiguous between a possessor and a pronoun. In the case of her ambiguity, permitted reanalysis does not implicate the rejection of Theta Attachment, because the argument structure of the theta assigner, the verb give, remained unchanged and maximal in its satisfaction of Theta Attachment, both in the initially-built representation and the repaired representation.

In addition to the lexical ambiguity of her, there are cases, not discussed earlier, where the cure, ‘permitted reanalysis’, involves selectional ambiguity of verbs.

(54) Ron knew the ugly little man hated him.

The verb know is locally ambiguous in its theta grid between a <THEME> and a <PROPOSITION>. It initially attaches the ugly little man as its direct object, discharging the <THEME> role in local satisfaction of the Theta Principle and Theta Attachment. When hated is encountered, a local violation of the Theta Principle is diagnosed: no argument is locally available for the <AGENT> role of hate to pMerge with. However, a cure is available. The cure involves reanalysis that does not violate the OLLC (lowering of the ugly little man):

```
Tree 12
```

```
TP NP TP
  Ron T VP
    V knew C TP
      the ugly little man T VP
        V NP
          hated him
```
The DP *the ugly little man* is lowered from the thematic (direct object) position of *know* to the subject thematic position of *hate*. The cure is permitted, since the source position governs the target position. The cure is also grammatically possible, since the verb *know* has the lexical property to select either an NP or a CP as its complement. There was no rejection of Theta Attachment because *know* has two realizations of argument structure, both are maximal: one with an NP (<THEME>) as complement, the other, a CP (<PROPOSITION>). ‘Permitted reanalysis’ is not costly because the cure does not involve the rejection of Theta Attachment.

Continuing with the flowchart, we consider what happens when the ambiguities so far described (categorical ambiguity and selectional ambiguity) are not involved.

Recall that there are cases of relicensing signalled by Adjacency (discussed in §1.4.3.2, and demonstrated in example (17) *The spaceship destroyed in the battle disintegrated*). When Adjacency is applicable (that is, when the verb is obligatorily transitive), it cancels any source or target position. The OLLC becomes irrelevant to cases of relicensing. If it is the case that Adjacency is relevant (viz. involving an obligatorily transitive verb), then the local violation of the Theta Principle is cured without cost.

If the cure does not involve the rejection of Theta Attachment (cases of ‘permitted reanalysis’) or Adjacency (viz. involve the special case of obligatorily transitive verbs), then the cure may involve the rejection of Theta Attachment (raising), if this cure is grammatically possible (when omission of a direct object is possible). It is in this latter case that the local violation of the Theta Principle is costly. The parser is punished for being guided by Theta Attachment: it had led the parser to a local violation of the Theta Principle. The rejection of Theta Attachment is something the parser is unable to do: it goes against its automatic mode of operation. However, it is required by grammar. The parser is punished for its adherence to Theta Attachment because Theta Attachment forced it to err, to locally violate grammar. A garden path effect is therefore the failure to avoid the local violation of the Theta Principle.

Finally, if the cures so far described are not grammatically possible, then the violation of the Theta Principle is not costly (in correspondence with Pritchett’s quote on page 43).

1.6.1. Internal restructuring & relicensing derive from Adjacency

It is possible to further simplify Flowchart 1 and the prediction of the garden path effect. Let us consider an example that involves internal restructuring (like example (15)).
(55) Mary sent her flowers to John.

The parse is conducted as usual, initially obtaining the full satisfaction of Theta Attachment: $[Agent Mary] sent [Goal her] [Theme flowers]$. When the real $<GOAL>$ to John appears, there is no theta-role available that may license it: the $<GOAL>$ role was pMerged with her. In order for the prepositional phrase not to remain unlicensed in local violation of Theta Attachment, internal restructuring becomes available as a cure given that her is lexically ambiguous.

Better local satisfaction of the Theta Principle can be achieved by having her flowers pMerged with the direct object position (and not flowers). Internal restructuring is required in (55) in order to fuse her and flowers by modification. This is the cure and it is permitted by the OLLC (the source position governs the target position as in example (15)). There is a grammatical alternative representation in which her is the possessor of flowers. Correspondingly, the violation of the Theta Principle is not costly.

In terms of Adjacency, the constituent her flowers is identified to be the one that needs to be right-adjacent to the verb. This constituent must occupy the direct object position of send. For this reason, the structure built thus far requires relicensing. The local violation of the Theta Principle is not costly because relicensing does not involve the rejection of Theta Attachment (the verb send retains its maximal theta grid in satisfaction of Theta Attachment).

Internal restructuring and relicensing (as in (17)) are two sides of the same coin: Adjacency, as a principle that expresses a relation between a direct object and a licensing element. In relicensing, the link to Adjacency is obvious: the absence of a direct object right-adjacent to a verb signals that the structure built so-far must be revised. In internal restructuring, the ‘real’ direct object that must be right-adjacent to the verb is identified, which is enough to cancel the applicability of the OLLC. Notice that in the case of permitted reanalysis (lowering), where a lexically ambiguous verb is involved (such as know in example (54)), Adjacency is also of relevance: the direct object that must be right-adjacent to the verb is not an NP but a CP (a CP is the direct object, not an NP).

We may conclude that Adjacency serves as a ‘repair’ signal, which circumvents the costly rejection of Theta Attachment. It tells the parser how the cure operation should be carried.

If it is grammatically possible to raise, then by necessity, the cure involves rejection of Theta Attachment and correspondingly the local violation of the Theta Principle is costly. Otherwise, it is not grammatically possible to locally satisfy the Theta Principle, which is inherently not a costly violation of the Theta Principle. The decision when the local violation of the Theta Principle is costly is summarized in Flowchart 2.
Flowchart 2: Deciding when the local violation of the Theta Principle is costly, conflated version

As can be seen, a costly local violation of the Theta Principle is manifested by a garden path effect, where the cure involves the rejection of Theta Attachment (raising, or violation of the OLLC). Adjacency involves relicensing and internal restructuring.

In the rest of this study, I will refer to the garden path effect as a local violation of the Theta Principle. It is used with the understanding that the local violation of the Theta Principle is costly only in those cases where the OLLC is violated (the cure is raising, involving the rejection of Theta Attachment).

Regarding the garden path effect as local ungrammaticality, i.e. the local violation of the Theta Principle, has the advantage that it enables the
generalization of the Thematic Transparent Model so formulated to other processing phenomena, such as center-embeddings. As I will show, these involve the local violation of the Theta Principle, too. By this, prominent processing phenomena are unified under a single cause and as a consequence there is better grounds for me to claim that the model corresponds with actual human NLP – it has reasonable ‘goodness of fit’. Further, I may claim with greater certainty that human parsing performance is guided by Theta Attachment.

1.7. Conclusion: a viable model

In this chapter, I discussed the ‘goodness of fit’ of the thematic parser. I presented the model, and brought about objections to its validity. I concluded that the Thematic Transparent Model was well motivated despite alleged counterevidence. The model’s ‘goodness of fit’, i.e. its conformity to given data and to external reality, has been retained in both left and right-branching languages.

I demonstrated under which circumstances the local violation of the Theta Principle is costly (when the OLLC is violated). Further, I showed that under such circumstances, the garden path effect is the local violation of the Theta Principle. The Theta Principle conforms to external reality as demonstrated by interruption in brain activity when a garden path construction is read.

I provided support for the validity of Theta Attachment as the guiding principle of initial resolution of attachment and for the parser’s reliance on thematic information from various fields of psycholinguistic research. Theta Attachment is a fundamental principle of the Thematic Transparent Model and as such, it is important to supply empirical evidence for it.

Theta Attachment makes a specific and unique prediction concerning the occurrence of processing difficulty in certain garden path constructions. This prediction distinguishes the thematic parser from other processing models. If the prediction turns out to fit in with actual performance, then it would serve as evidence that human sentence parsing is motivated by Theta Attachment and that the thematic parser captures actual performance. The following chapter spells out this prediction and provides experimental evidence that indeed the prediction is compatible with actual parsing performance. The chapter serves as evidence that human sentence processing is guided by Theta Attachment.
2. THETA ATTACHMENT

2.1. Cases of Theta Attachment indeterminability

The former chapter has established the thematic parser as a viable model of human parsing performance. In this chapter, I will concentrate on the model’s ‘goodness of fit’ with actual human performance. In Section 2.1, I will present a specific prediction of the Thematic Transparent Model. In Section 2.2, I will discuss other processing models and their own predictions. The discussion will lead to the conclusion that the thematic parser is distinguished from other models with regard to its prediction in these specific cases, as is summarized briefly in Sections 2.3 and 2.4. Sections 2.5-2.7 provide the experiments that I have conducted and Section 2.8 concludes that the predictions of the thematic parser and only its predictions are compatible with the obtained empirical evidence.

2.1.1. Cases of theta-role surplus

Theta Attachment is a central principle of the thematic parser. Theta Attachment guides the initial resolution of local ambiguity during parsing. There are situations where Theta Attachment does not adjudicate: there are two equal alternatives in terms of the satisfaction of Theta Principle. This unique situation is demonstrated by example (56):
(56) ÷A waitress poured the visitors who were drinking red wine.\textsuperscript{20}

At the point where \textit{red wine} is identified, the \texttt{<THEME>} role of \textit{pour} is still lacking a nominal argument with which it may pMerge. At the same time, the \texttt{<PATIENT>} role of \textit{drink} also misses a nominal argument. A single argument is available for two theta-roles: \texttt{<THEME>} from \textit{pour}, and \texttt{<PATIENT>} from \textit{drink} – this is a situation of theta-role surplus.

Theta Attachment by itself does not force one choice over the other, since either assignment will provide \textit{red wine} with a theta-role, releasing a single role from one grid and stranding another and thus satisfying the principle equally.

If the parser pMerges \textit{red wine} with the \texttt{<THEME>} theta-role of \textit{give}, a grammatical structure will be built and the Theta Principle will be satisfied. If \textit{red wine} pMerges with the \texttt{<PATIENT>} theta-role of \textit{drink}, then there will be a thematic position without an argument NP (the \texttt{<THEME>} role of \textit{pour}). The cure entails the rejection of Theta Attachment (therefore violating the OLLC, too, since the cure includes the illicit raising of \textit{red wine}), and therefore involves a costly local violation of the Theta Principle:

\textsuperscript{20}The sign ÷ denotes a garden path construction with thematic indeterminability.
Local ambiguity in this case has two resolutions consistent with Theta Attachment: one which is predicted to be easy, and another which leads to processing difficulty. Since Theta Attachment has no preference for one attachment choice over the other, the resolution of the ambiguity will be at random.

2.1.2. Non theta-role surplus cases

Other indeterminable situations involve two options that equally satisfy the Theta Principle but there is no theta-role surplus. Such examples involve verbs locally ambiguous between a passive participle and a transitive theta grid:

Tree 13: illicit reanalysis in theta-role surplus sentences

As Pritchett points out in his discussion of these structures (§3.3.2 and §4.0.4.2 of his 1992 book), global obligatoriness of the theta-roles is locally irrelevant for Theta Attachment.
The child just found ran away.

*Found* is locally ambiguous between a passive participle (having a `<THEME>` role) and transitive theta grid (note that *find* lacks an intransitive entry, like *race* in (16)).

(i) Should the parser choose the transitive entry, only the `<AGENT>` theta-role would be locally discharged but not the `<THEME>` role.

(ii) If the parser were to choose the passive participle interpretation, then the `<THEME>` role would be discharged on an empty category (a gap) indexed with a (null) operator, in a relative construction.

Although a single thematic role is discharged in the passive and transitive analyses; in the transitive analysis the `<THEME>` awaits discharge and in the passive analysis there is a relative NP awaiting a theta-role.

This introduces a situation where both options can satisfy Theta Attachment to the same degree. Should the passive participle option be chosen, the relative NP *the child* will remain unattached. If the transitive option is pursued, there is no argument to which the internal theta-role can be attached. The transitive option will lead to a processing difficulty, since *ran away* will lack an argument to pMerge with its external thematic position. This is a costly local violation of the Theta Principle. If the passive participle option is chosen, then the correct interpretation of the sentence is built, and no violation will occur. Accordingly, attachment is predicted to be resolved at random. One option would lead to unprocessability, the other to grammaticality.

Sentence (58) is a similar case, involving permitted reanalysis. The verb *know* may select a noun argument or a sentential complement. The noun argument interpretation is chosen since the guy can be readily attached in local satisfaction of Theta Attachment. When *hurry* enters the parse, it requires an argument NP to pMerge with its external thematic position.

(58) Mary knew the guy hurried into marriage would divorce.

However, the verb *hurry* is locally ambiguous between three thematic entries from the lexicon: (a) intransitive, (as in *the guy hurried*); (b) transitive with an optional `<GOAL>` role (as in *the guy hurried the girl into marriage*), and (c) the passive participle. Option (b) strands the `<THEME>` role (and the `<GOAL>` role), and is therefore not pursued. Options (a) and (c) are equal in their satisfaction of Theta Attachment. In option (a), the DP *the guy* receives a theta-role from *hurry*, and the Theta Principle is satisfied. In option (c), the `<THEME>` role is discharged on an empty category creating a relative NP. Since *the guy* has already received a theta-role from *know*, the relative NP does not await a theta-role, and this option satisfies the Theta Principle, too.

If option (a) is chosen, this will inevitably lead to a local violation of the Theta Principle. When *divorce* is entered, its (external) thematic position
Theta Attachment

will be deprived of an argument NP for pMerge. The situation renders the local violation of the Theta Principle costly. However, if option (c) is pursued, then a grammatical representation that satisfies the Theta Principle will be yielded, predicting no processing difficulty. Attachment is resolved at random yielding in one case processing difficulty, in the other its absence.

The prediction of random performance on sentences with Theta Attachment indeterminability is unique to the Thematic Transparent Model and separates it from other processing models of garden path constructions. Should evidence be found to support this prediction, it will serve to show that Theta Attachment guides human sentence parsing. It is therefore important to establish whether the prediction fits with actual performance.

In Section 2.2, I will present other prominent processing models of garden path constructions, and their specific predictions with regard to theta-role surplus type of sentences.

2.2. Predictions of non Theta Attachment models

2.2.1. The Garden Path Model

The Garden Path Model is different from the Thematic Transparent Model, since it does not consider lexical information as a guiding principle in the resolution of initial syntactic attachments. This echoes the long-standing debate described in §1.5.1.4, whether thematic information guides the parser in making attachments decisions or not. Rather, the model relies on complexity estimated by the number of nodes as represented by the syntactic tree, preferring the less complex attachment (i.e. the attachment with the least number of nodes). Two principles are central for the Garden Path Model:

(59) **Minimal Attachment**: Do not postulate any potentially unnecessary nodes.

(60) **Late Closure**: If grammatically permissible, attach new items into the clause or phrase currently being processed (i.e. the phrase of clause postulated most recently).

In the parsing of sentence *The horse raced past the barn fell*, e.g., readers incorporate each word of an input into a constituent structure representation of the sentence, roughly as each item is encountered as illustrated in (61) (example taken from Frazier (1987a)). At each step, the readers postulate the

---

22 The PP *into marriage* does not signal relicensing as constrained by Adjacency, since the sentence *Mary knew the guy hurried into marriage* is fine, given the intransitive theta grid.
minimal number of nodes required by the grammar of the language under analysis, given the structure assigned to processing items.

(61) Parsing steps according to the Garden Path Model

a. b. c. d. e.

Minimal Attachment predicts that the transition from (61d) to (61e) will require revision of analysis, i.e. be difficult, since a relative NP requires more nodes.

Late Closure adjudicates in cases where two equally Minimal Attachments exist. It will favor attachments to preceding items (over attachment to subsequent items) and typically will favor attachments to phrases lower in the phrase-structure tree rather than to phrases higher up.

In the case of the theta-role surplus sentence (56), repeated here as (62), no random performance is expected. At the point red wine is processed, Minimal Attachment does not adjudicate, since the low attachment of red wine (to drink) is equal in terms of nodes to its high attachment (to pour).

(62) A waitress poured the visitors who were drinking red wine.

Late Closure resolves this situation by preferring the low attachment: the phrase postulated most recently is that of drink, and it is grammatically
possible to attach red wine low. Revision of analysis is required to obtain a grammatical representation and a garden path effect is yielded. The locally ambiguous situation (when Minimal Attachment does not adjudicate), is always resolved by Late Closure, and therefore the Garden Path Model does not predict this situation to be resolved at random. The model predicts that theta-role surplus sentences will always incur a garden path effect (= processing difficulty).

2.2.2. The Diagnosis Model

The Diagnosis Model (Fodor and Inoue, 1994, 1998) is a repair-based model. By ‘repair’ is meant that the parser decides what structural alterations to make when a wrong analysis of the input has been detected, but it does not reparse the input.

To illustrate the model, consider the following garden path sentence containing object-subject ambiguity:

(63) While the boy scratched the dog yawned loudly.

The guiding principle of the parser is Attach:

(64) Attach: On receiving a word of the input sentence, connect it to the Current Partial Phrase Marker (CPPM) for the sentence in such a way that the resulting CPPM is syntactically well-formed though possibly incomplete at its right edge. (Fodor and Inoue, 1998, p. 103)

Up until yawned, the parser creates a CPPM that is syntactically well-formed While the boy scratched the dog. When the input word yawned enters the CPPM, it is attached anyway, although this violates grammar:

(65) Attach Anyway: Having established that there is no legitimate attachment site in the CPPM for the current input word, attach the input into the CPPM wherever it least severely violates the grammar, and subject to the usual preference principles that govern Attach. (ibid, p. 105)

The temporary ill-formedness within the CPPM is immediately resolved by the parser under the guidance of Adjust:
Adjust: When a grammatical conflict has been created between two nodes or features X and Y in the CPPM, by either Attach Anyway or Adjust, eliminate the problem by altering minimally (i.e. no more than is necessary for conflict resolution) whichever of X and Y was less recently acted on, without regard for grammatical conflicts thereby created between that node and other elements in the CPPM. (ibid, p. 106)

The illegitimate attachment converts a mismatch between the CPPM and subsequent input into a mismatch internal to the CPPM. At that point, “grammar can be called on to identify the nature of the problem” (ibid, p. 107). Repair is difficult in case the parser cannot easily deduce at each step of the correction what the nature of the conflict is and how to cure it. The Adjust operation is constrained by the Grammatical Dependency Principle:

Grammatical Dependency Principle (GDP): When a grammar violation has been created in the CPPM by an action on node n in accord with Attach Anyway or Adjust, attempt to eliminate the problem by acting on a node that is grammatically incompatible with n. (ibid, p. 109).

In (63), the symptom is lack of subject for yawned. Since there is no grammatical dependency that links the matrix subject position where the grammatical problem is, to the object position in the prepositional adjunct, where the repair has to be performed, revision is difficult. GDP cannot guide the repair operation and it is therefore difficult.

The Attach principle within the Diagnosis Model involves the same principles of the Garden Path Model, namely Minimal Attachment and Late Closure. With regard to theta-role surplus sentences, such as (62), the prediction of the Diagnosis Model would be the same as the Garden Path Model’s prediction. A garden path in these sentences is always predicted. It is only that the source of the difficulty is explained differently: there is no grammatical dependency that links the direct object position of drink where the grammatical problem is, to the direct object position of pour, where the repair has to be performed, resulting in difficulty.

2.2.3. The Snip Model

Lewis (1993) suggests that parsing consists of ‘Link’ and ‘Snip’ operations. The parser creates nodes by the Link operation. When the input is incompatible with the structure built so far, the Link operator attempts to attach it to the current phrase, even though it temporarily creates an illegitimate syntactic structure. The repair operation Snip detaches from the tree built so far the violating node. The Snip operation can only be initiated
locally, i.e. within the minimal maximal projection containing the inconsistency.

In (68), *food* is initially linked to the direct object position of the VP headed by *eat*. When *is thrown* is encountered, there is a grammatical inconsistency because *is thrown* misses its obligatory subject.

(68) ¿After Fred eats food is thrown.

The correct repair operation would be to cut *food* from the current position and re-link it to the subject position of the newly formed TP. However, the Snip operation cannot be initiated because the position at which the grammatical inconsistency is detected is not local to the position at which Snip would have applied (*food* is within a VP but not the TP). The structure, therefore, cannot be repaired, and is difficult to process.

Now let us consider the theta-role surplus sentence (69).

(69) A waitress poured the visitors who were drinking red wine.

As this model is constructed, a problem arises only when the Snip operation cannot be initiated. The model assumes that the Link operation is local, and therefore prefers the initial low attachment of *red wine*. The required Snip operation will have to be outside of the VP headed by *drink*, and therefore costly. No theta-role surplus is predicted, and a garden path effect is always predicted for this type of sentences.

2.2.4. The LAST Model

The Late Assignment of Syntax Theory (LAST) suggests that the parser is initially driven by the recognition of the most statistically dominant thematic patterns (Townsend and Bever, 2001). The most statistically dominant thematic pattern is NVN = [agent action patient]. This surface sequence provides reliable information for initial syntactic analysis. Consider the following object-subject ambiguity example:

(70) ¿While Mary was mending the sock fell off her lap.

The parser (system, in their terms) matches the NVN pattern to *Mary was mending the sock* and assigns *Mary* as the agent, *was mending* as the action, and *the sock* as the patient. Based on this initial meaning-form hypothesis, a separate comprehension system generates a syntactic structure, and checks it against the input sequence. Since the candidate syntactic structure matches the input sequence, this meaning is stored. When *fell* is received, the system cannot match an initial verb with the NVN pattern and an error is detected. The assignment of *the sock* as <PATIENT> of *was mending* is replaced by accessing a different subcategorization of the verb (as an intransitive) and reassigning *the sock* as <THEME> of *fell*. 
A garden path is expected in theta-role surplus sentences at all times according to the LAST Model.

(71) A waitress poured the visitors who were drinking red wine.

The sentence incurs difficulty because of the interaction of subcategorization requirements for pour and drink. The strength of the NVN template leads the system to propose that red wine is <PATIENT> of drink. Assigning red wine as <PATIENT> of drink leaves unfilled the third argument of pour, which has the pattern typical of ditransitive verbs: NVNN = [agent action recipient patient]. These patterns interact with one another and a garden path is predicted at all times due to the system’s inability to decide which pattern fits best.

It should be noted that the LAST Model discusses the question whether a garden path effect will be produced or not in a structure similar to (71). In (72), the sentential complement to plow the field can be attached to either hired or told:

(72) Ian told the man that he hired [to plow the field] [a story].

Townsend and Bever discuss the question whether (72) produces a garden path. The LAST Model is indecisive about its prediction of the garden path effect – that it, it seems that the model itself cannot make any specific prediction whether there is a garden path or not. They conjecture that: “[i]t seems likely that intonation in normal speech will prevent a garden path” (Townsend and Bever, 2001, p. 316). They do indicate that the sentence raises the interesting question of the extent to which subcategorization information is available to guide the system’s development of an initial meaning-form hypothesis, and to the extent to which the system relies on the NVN template. If the system relies strongly on the fact that tell requires either an additional noun phrase argument or a sentential complement, they expect that to plow the field will complete a semantic unit for tell. In this case, the system will have to override the NVN template in the relative clause, and present the meaning that the patient of hire is man. Since this analysis requires undoing the NVN template, such sentences provide an interesting test of the relative importance of an independent NVN template and subcategorization information.

2.2.5. Structural Determinism

In models based on structural determinism (Gorrell, 1995a, b, 1998, and Sturt and Crocker, 1996), primary structural relations, dominance and precedence (but not immediate domination and precedence) take central stage.
(73) **Structural Determinism**: the domain of determinism is limited to the primary structural relations, dominance and precedence.

Once these relations are determined, they may not be tampered with. Other structural relations, such as government and theta-role licensing are considered secondary relations and can be altered during parsing. The only additions consistent with Structural Determinism are equivalent to “inserting one tree inside another at an immediate point on the right frontier” (Sturt and Crocker, 1996, p. 451).

In the now familiar object-subject ambiguity garden path (74), Structural Determinism is violated during parsing.

(74) ¿After Fred eats food is thrown.

When *food* enters the parse, it is assumed to be within the VP headed by the verb *eat*. The primary relations so far established are given in (75).

(75)

When *is thrown* is encountered, a new TP is added. *Food* has to be the subject now. The relation between VP1 and NP2 has to be erased:
The parser may not erase the initial dominance and precedence relations for NP₂, because it is strictly deterministic. A garden path effect is the result.

The prediction of Structural Determinism for cases of theta-role surplus is quite straightforward, since it relies on attachment to the ‘right frontier’, similar to Late Closure in the Garden Path Model. The NP red wine will always be attached to the right frontier: the lower verb VP₂.

(77) A waitress [VP₁ poured the visitors who [VP₂ were drinking red wine]].

The domination of red wine by the lower verb VP₂ would have to be erased in favor of its domination by the upper verb VP₁, producing the garden path effect.
(78) A waitress [vp_1 poured the visitors who [vp_2 were drinking] red wine].
There is therefore no question of indeterminability as expressed by Theta Attachment.

2.2.6. Memory-Based Ranked Parallel Model

Gibson (1991) proposes a parallel parser, which evaluates alternative analyses during parsing. These alternatives are ranked in terms of memory cost, when each new word is received. Any analysis with a higher memory cost than its competitor, or with a memory cost which exceeds an absolute limit on computability, is discarded. The cost of maintaining a particular analysis in memory is evaluated by the number of violations of certain general constraints such as the Theta Criterion within Government and Binding Theory (Chomsky, 1981, 1986), and cognitive resources such as recency effects. Memory load is estimated by Processing Load Units (PLUs):

(79) The property of thematic reception (TR): Associate a load of $X_{TR}$ PLUs to each C-node (confirmed node) constituent that is in a position that can receive a thematic role in some co-existing structure, but whose theta-assigner is not unambiguously identifiable in the structure in question.

(80) The property of lexical requirement (LR): Associate a load of $X_{LR}$ to each lexical requirement that is obligatory in some co-existing structure, but is satisfied by an H-node (hypothesized node) constituent containing no thematic elements in the structure in question.

Gibson notes that $X_{TR}$ and $X_{LR}$ are approximately equal, so both can be referred to as $X_{int}$, where ‘int’ stands for ‘interpretation’.

In the parsing of object-subject ambiguity sentences, such as (81), there is no cost at the point of processing food.

(81) ¿When Fred eats food is thrown.

Analysis (82a) given below is obtained since there are no NPs in theta positions without an identified theta assigner, or elements with dissatisfied lexical requirements. However, the string in (82) could have the structure in (82b) and could turn out to continue as in (81), which causes processing difficulty.

(82) When Fred eats food…
   a. [When Fred eats food]…
   b. [When Fred eats] [food]…
For analysis (82b), at the point of processing *food*, the cost is $2X_{int}$ PLU. $1X_{int}$ PLU is due because a theta assigner for *food* has not been identified (principle (79)). An additional $1X_{int}$ PLU is counted because the lexical requirement for the NP *food* to receive a theta-role is not satisfied. The omission of the <THEME> role of *eat* is not costly since it is not obligatory (as principle (80) dictates).

The difference of $2X_{int}$ PLU is enough for the parser to abandon the correct structure (82b) for (82). A severe garden path effect is predicted at the disambiguation point (when the final verb *is thrown* is encountered) – this alternative analysis is ranked lower.

What happens in cases of theta-role surplus according to the memory-based ranked parallel model?

(83) A waitress poured the visitors who were drinking red wine.

At the point of parsing *red wine*, two theta-roles are available for *red wine*. The ranked model solves the deadlock by incorporating recency effects, essentially preferring the local (linear) attachment of *red wine* to *drink*.

The transfer to the alternative correct analysis (high attachment of *red wine*) that must be built is not costly in terms of memory PLUs. According to (80), there is additional $1X_{int}$ PLU associated with not satisfying the lexical requirement of *pour* to discharge its obligatory <THEME> theta-role. $1X_{int}$ PLU is insufficient to incur difficulty ($2X_{int}$ PLUs are minimum). There is virtually no cost associated with the transfer from the low attachment analysis to the high attachment analysis ranked lower. It is therefore the prediction of this model that no difficulty will be reported in theta-role surplus sentences.

### 2.3. Two opposing groups of prediction

We may summarize the prediction of the various models with regard to performance on theta-role surplus sentences.

<table>
<thead>
<tr>
<th>The model</th>
<th>Occurrence of the garden path effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Thematic Parser</td>
<td>random</td>
</tr>
<tr>
<td>The Garden Path Model</td>
<td>always</td>
</tr>
<tr>
<td>The Diagnosis Model</td>
<td>always</td>
</tr>
<tr>
<td>The Snip Model</td>
<td>always</td>
</tr>
<tr>
<td>The LAST Model</td>
<td>always</td>
</tr>
<tr>
<td>Structural Determinism</td>
<td>always</td>
</tr>
<tr>
<td>Memory-Based Ranked Parallel Model</td>
<td>no effect</td>
</tr>
</tbody>
</table>

Only the Thematic Transparent Model predicts performance at random on theta-role surplus sentences, and it is therefore set apart from other models.
All other models, except the Memory-Based Ranked Parallel Model, predict a consistent occurrence of the garden path effect. The Memory-Based Ranked Parallel Model predicts no occurrence of the garden path effect.

The prediction of random occurrence implies also another type of prediction. All models, except the Memory-Based Ranked Parallel Model, predict performance on cases of thematic role surplus and other types of garden paths (say, the canonical example “The horse raced past the barn fell”) will not differ. Both types of constructions will manifest processing difficulty. The Memory-Based Ranked Parallel Model predicts that performance on theta-role surplus sentences would not differ from sentences that does not incur processing difficulty. Only the Thematic Transparent Model predicts that theta-role surplus sentences will differ in performance from other canonical garden paths and from non-garden path sentences. This is summarized in the following table.

<table>
<thead>
<tr>
<th>The model</th>
<th>Similarity of performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Thematic Parser</td>
<td>In-between canonical garden paths and unproblematic sentences</td>
</tr>
<tr>
<td>The Garden Path Model</td>
<td>Like canonical garden paths</td>
</tr>
<tr>
<td>The Diagnosis Model</td>
<td>Like canonical garden paths</td>
</tr>
<tr>
<td>The Snip Model</td>
<td>Like canonical garden paths</td>
</tr>
<tr>
<td>The LAST Model</td>
<td>Like canonical garden paths</td>
</tr>
<tr>
<td>Structural Determinism</td>
<td>Like canonical garden paths</td>
</tr>
<tr>
<td>Memory-Based Ranked Parallel Model</td>
<td>Like unproblematic sentences</td>
</tr>
</tbody>
</table>

This type of prediction concerns the means of performance over a pool of subjects. If performance on theta-role surplus sentences is at random (either a garden path effect or not, 50:50 chance of a garden path effect), then this type of performance will be reflected by a performance measure (a mean) that will be in-between canonical garden paths and sentences without processing difficulty, but one which will be different from both.

In these respects, the Thematic Transparent Model stands out in stark contrast to other processing models. If empirical support could be found that performance is at random, then it will serve as support for Theta Attachment as a guiding principle of initial resolution of ambiguity in human sentence parsing. Furthermore, if performance on theta-role surplus sentences were different from performance on canonical garden paths and non-garden paths, then this would also serve as evidence for the predictions of the thematic parser.
2.4. Preliminary empirical evidence for random performance on theta-role surplus sentences

Pritchett, relying on an informal survey, notes that the prediction concerning the performance on theta-role surplus constructions as in (56) was borne out by empirical data: individuals were inconsistent with regard to tokens of such sentences judging them as either unprocessable or unproblematic. This is consistent with the hypothesis that in instances where there are two (or several) equal attachment possibilities, the situation is resolved essentially randomly.

No further information is supplied concerning the surveys Pritchett conducted. However, there are indications in the literature of ‘lack of preference’ for one interpretation over the other in constructions that involve final PP ambiguity. Lack of preference is typically indicated by 50:50 performance (around 50% of the subjects indicate low attachment). Ford, Bresnan and Kaplan (1982) report no bias in subjects’ interpretation of sentences involving final PP ambiguities:23

(84) Sue discussed the difficulties that her daughter was having with the teachers.

The final PP with the teachers can be attached either to have or to discuss. Sixty percent of 20 subjects tested by Ford, Bresnan and Kaplan (1982) understood the sentence to mean her daughter was having difficulties with the teachers.24 The rest said the interpretation was that Sue discussed them (the difficulties) with the teachers.

Notice that (84) is globally ambiguous and is not a garden path. This means that the question whether subjects experience processing difficulty essentially at random in (56) cannot be deduced from (84): these are two different sentences in terms of performance. Example (56) is locally ambiguous, (84) globally. What is interesting about (84) is that comprehenders (as a group, but perhaps not as individuals) demonstrate that they are aware of the grammatical possibilities of PP attachments, therefore applying grammatical competence to sentence interpretation.

Sadeh-Leicht (2003a, 2003b) tested theta-role surplus sentences in Hebrew in the purpose of establishing whether reports of garden path effects were at random. The sentences were similar to sentence (56):

---

23 The study was a multiple-choice questionnaire containing the ambiguous sentences plus distracters. The subjects had to decide which choice was a possible interpretation.

24 A result of 60% is not significantly different from the predicted 50:50.
In (85), the NP apple may be attached to ate or to peeled. Attaching apple to ate results in a costly violation of the Theta Principle. If it is attached to peeled, the grammatical representation is obtained.

In this off-line questionnaire study, subjects were asked to read tokens identical to (85). They had to compare the test tokens with two ‘reference’ sentences. One was a garden path sentence (object-subject ambiguity), the other was a non garden path. The subjects were instructed to decide whether tokens of theta-role surplus were difficult to process or not. The prediction was that only 50% of the subjects would say that (85) is comparable in difficulty to the object-subject reference sentence.

There were 106 participants in this study. 64% of the subjects indicated difficulty on tokens such as (85). Although this result is not the desired exact 50%, it is statistically not different from random report of processing difficulty. For object-subject ambiguity (garden path) sentences, 95% of the subjects indicated difficulty (in the control unproblematic sentences, only 2% of the subjects reported difficulty). This serves to show that performance on theta-role surplus sentences was significantly different from performance on other garden paths and non-garden paths, and was at random. However, the Hebrew questionnaire is an off-line study and may not reflect actual human performance during real-time parsing.

Viewing processing difficulty in garden path constructions as local ungrammaticality (see Section 1.6) gave rise to the idea that techniques used for measurement of grammaticality judgment could be harnessed in resolution of the problems described in the previous paragraph. My choice fell on an experimental method called Magnitude Estimation, belonging to the developing field of experimental syntax (Cowart, 1997). Since Magnitude Estimation is a developing technique within linguistics, several experiments were necessary to make sure it is valid and capable of answering the relevant questions about performance on theta-role surplus sentences. Furthermore, it was necessary to establish whether this off-line technique corresponds with on-line techniques.

Experiment 1 in Section 2.5 (on English) establishes whether this off-line method corresponds with an on-line procedure, such as eye-tracking.

Experiment 2 in Section 2.6 (on English) validates Magnitude Estimation as a suitable tool for measurement of the garden path effect as defined within the thematic parser.
Experiment 3 in Section 2.7 is concerned with the question whether theta-role surplus sentences in Dutch yield different performance than that of other standard garden path constructions, and whether the occurrence of the garden path effect is actually at random.

2.5. Experiment 1: Validation of Magnitude Estimation

2.5.1. Introduction

Magnitude estimation is an experimental method developed within the field of psychophysics to measure sensory thresholds of subjective sensations of physical stimuli (light, sound, etc.) in a reliable fashion (Stevens, 1951).

In a classical psychophysical experiment, subjects are presented with a series of sensory stimuli (say, varying light intensities) one at a time in random order across a wide range of stimulation and are instructed to give numbers to the perceived brightness of each stimuli relative to the first (light intensity), sometimes referred to as modulus. If a given light seemed 2 times brighter than the first light (modulus), the subject would give a number twice as large as the modulus; if one-third as bright, a number one-third as large and so on for a range of light intensities. Characteristically to the magnitude estimation procedure is that all judgments are made explicitly relative to a reference. The estimations can be plotted against the empirical scale of light intensities (obtained for example from a machine that measures light intensity). It is then possible to establish the exact subjective threshold of light intensity sensed by the subject.

The use of Magnitude Estimation in psychophysics and subsequently in social sciences is due to several advantages the method offers over category scaling in the measurement of psychological dimensions (in category scaling a fixed number of options is selected: for instance, a variety of 2-, 3-, 5-, 7-, and 9-point category scales). One disadvantage of category scaling is that information is lost because of the limited resolution of the categories. Category scaling forces subjects to make similarity-difference judgments and place items judged as ‘more’ or ‘less’ alike into the same category. The greater the discrepancy between the true range of the stimuli and the fixed range of the category scale, the poorer the resolution of the categories. For this reason, categorical judgments may be thought of as qualitative rather than quantitative. Second, category scales represent only an ordinal level of measurement, thereby denying researchers legitimate access to many of the powerful parametric statistical methods based on interval assumptions, which are available today. Third, by offering a fixed number

---

25 An ordinal scale has two formal properties: equivalence and order. If two objects are the same with
of categories, however few or numerous, the researcher is inadvertently affecting the response. Category judgments may be constrained, at times providing a too expansive range and only on occasion matching fixed to true range. Furthermore, it was found over the years that there was a low to moderate degree of correspondence between category scaling and behavior (Lodge, 1981).

Although Magnitude Estimation has been suggested a long time ago to be potentially useful for linguistic research (see Chomsky, 1961), it has only been recently adopted for syntactic investigations. The major problem was that Magnitude Estimation required a comparison of subjective sensations to an objective physical scale. In experiments designed to test the threshold of human beings’ perception of hue change, for example, differences in color stimuli could be measured by an instrument and then estimations could be compared against the empirical measurements. In linguistics (and social sciences), such an objective scale seemed unattainable (Schütze, 1996). However, Bard, Robertson and Sorace (1996, and Sorace and Keller, 2005) suggested a solution to this problem: cross-modality matching, following what has been suggested for social sciences, as in Lodge (1981).

In the psychophysical version of cross-modality matching, subjects use one sensory modality to estimate magnitudes presented in another. For instance, brightness might be estimated by adjusting the length of a line to correspond to the perceived brightness of a light. If the subject thinks that the second light stimulus is twice as bright as the first, he or she draws a line that appears twice as long as the line drawn to represent the brightness of the first stimulus. The plot of judgment (line length) against stimulus (light energy) was found to be characterizable by a straight line (in log-log scales). The logic behind cross-modality matching was that if subjects were shown to match in judgments of brightness and line length, then subjects can judge proportions of stimuli differences in two different modalities.

Cross-modality matching was validated for grammaticality judgments by Bard, Robertson and Sorace (1996). They showed that subjects were able to judge the grammaticality of sentences by using line lengths. This outcome is important since it enables the experimenter to know whether subjects internalized the concept of proportional estimations of line-lengths and understood the instructions given to them. The experimenter may be confident that subjects have also applied this concept to judgments of sentence acceptability.
What type of sensation is judgment of acceptability? Two dimensions can be defined in sensation. A prothetic dimension deals with sensations that differ in quantity. A metathetic dimension deals with sensations that differ in quality. In the prothetic dimension, a change in a particular stimulus results in a change in the intensity of the perception of that stimulus, that is, it is an additive change (bigger/smaller, brighter/darker). In the metathetic dimension, a change in particular stimulus results in a non-additive change in perception. For visual stimuli then, brightness (which can be darker or brighter) is in the prothetic dimension and hue (which can be red or green) is in the metathetic dimension.

Changes in judgments of local grammaticality are additive, therefore prothetic. Garden path constructions are locally ungrammatical within the Thematic Transparent Model. Hence, my aim was to reliably measure additive differences in local judgments of grammaticality.

For practical reasons, it is not possible to instruct subjects to judge local grammaticality. However, subjects do have an intuitive notion of complexity. Garden path constructions are complex by default, viz. local judgment of grammaticality equates with complexity. Subjects were therefore instructed to estimate complexity of garden path constructions, which was equated with local judgments of grammaticality.

In the linguistic application of this method provided here, subjects are asked to assess their personal sensation of complexity of a reference sentence (the modulus) by assigning it a number of their choosing. Subsequently, they are asked to estimate the complexity of experimental sentences in proportion to the modulus. Sentences that are more complex should be estimated higher than the modulus; less complex sentences should be estimated lower than the modulus. Magnitude Estimation allows free choice of numbers, as mentioned, and therefore its advantage is that it more accurately reflects actual sensations and subtle differences between them than standardly used category scales (such as, the semantic differential or the Hebrew questionnaire described in Section §2.4). Another version (not employed here) is to give subjects an arbitrary number said to reflect the complexity of the modulus, and let subjects estimate stimuli in relation to that arbitrary number.

Note that any quantity may be used for Magnitude Estimations, whether it is complexity, comprehensibility, plausibility etc., because estimation uses a relative scale (ratio), lacking units of measurement. In garden path constructions, the relation can only be associated with a sensation of perturbation, which may be referred to as ‘complexity’, ‘processing difficulty’, ‘comprehensibility’, or any other theoretical quantity of choice. It is the theory that determines the quantity.
Another question concerning Magnitude Estimation, being an offline method, is whether it is suitable to measure real-time processing phenomena, such as the garden path effect. The method has been used as a tool in the study of processing (Keller, 2003) However, it has rarely, if at all, been applied to the measurement of processing difficulty, a real-time phenomenon, which is a result of garden path constructions.

My initial aim was validation of this off-line technique for the studies of on-line processing. This requires comparing Magnitude Estimation with a well-established on-line technique. Eye-tracking is said to reflect actual performance on garden path sentences (Rayner, 1998, Rayner and Sereno, 1994). To this end, sentences from an eye-movement experiment by Frazier and Rayner (1982) concerned with processing difficulty were taken and used for validation of the Magnitude Estimation method.

2.5.1.1. Background to the eye movement experiment

Frazier and Rayner (1982) devised an experiment to record subjects’ eye movements while reading sentences with temporary structural ambiguities. They were interested in finding out the strategies subjects employed to repair such ambiguities. However, in order to test this, they also had to establish that processing difficulty occurred in local structurally ambiguous constructions, and that this difficulty was reflected in eye-movements (higher number of fixations or longer fixation durations in a certain region or overall). They indeed found longer reading times for sentences with local structural ambiguity compared to unambiguous controls. Therefore, their experiment was suitable for establishing the validity of the Magnitude Estimation procedure.

Sentences such as (86) were used in Frazier and Rayner’s experiment.

Non-Minimal Attachment, Short

(86) His second wife will claim the inheritance belongs to her.

According to the Garden Path Model, readers are misled to interpret the temporarily ambiguous noun phrase the inheritance as the direct object of the verb claim, but subsequently find out it is the subject of the complement clause. This violates Minimal Attachment (an S-node is required) and therefore processing difficulty is experienced.

Difficulty is not expected to occur if the noun phrase is indeed the direct object of the verb, as in (87):

Minimal Attachment, Short

(87) The lawyer thinks his second wife will claim the inheritance.

Frazier and Rayner’s (F&R henceforth) purpose was to compare (86) and (87). If (86) were read slower (on the locally ambiguous noun or after it due
to spill-over effects) than (87), then this would have been taken as evidence for a garden path effect.

It is important to note that the definition of the garden path effect differs between the Garden Path Model and the Thematic Transparent Model – a source of much confusion. According to the Thematic Transparent Model, a garden path effect involves local ambiguity and severe processing difficulty. The Garden Path Model defines a garden path effect as any type of confusion with structural implications.

Sentence (86) involves permitted reanalysis according to the thematic parser. This is reflected by a mild processing effect. Therefore the sentence would not be classified as demonstrating a garden path effect (it also does not involve a costly locally violate the Theta Principle). Within the Garden Path Model, any type of erroneous structural attachment on the part of the parser results in difficulty without making a distinction in the degrees of processing difficulty or whether reanalysis is conscious. The sentence is classified as a garden path within this model. Accordingly, the experiment described in this section validates Magnitude Estimation as a method that is sensitive enough to sense Minimal Attachment violations or ‘permitted reanalysis’ effects (but not a conscious garden path effect which is only found in the thematic parser and has a special status for this model only).

Continuing with the description of the eye-tracking experiment, F&R controlled for the length of the ambiguous NP. The reason is not indicated. Perhaps the lengthening of the ambiguous NP was required because of the method applied (measurement of eye-movements): If an ambiguous region is longer, then the probability of measuring an effect in that region will increase. No specific prediction was given in F&R’s experiment how the length of the ambiguous noun phrase could affect processing difficulty, though. But still this was helpful in testing the sensitivity of Magnitude Estimation, since length is known to affect judgements (van Kleeck, 1982).

F&R lengthened the ambiguous noun phrase, see sentence (88). They predicted that (88) should demonstrate processing difficulty, but (89) would be easier in relation to (88).

Non-Minimal Attachment, Long
(88) His second wife will claim the entire family inheritance belongs to her.

Minimal Attachment, Long
(89) The lawyers think his second wife will claim the entire family inheritance.

F&R indeed found that ambiguous versions of (86) and (88) overall took longer to read than unambiguous versions of (87) and (89). The results are
depicted in Figure 1 below relying on the total reading time per character.\footnote{Total reading time per character: the division of total reading time per sentence by the number of character spaces (including spaces between words).} There was an overall significant effect of Minimal Attachment; however, there was no overall effect of length. There was a significant interaction between length and Minimal Attachment.

![Figure 1: Reading time per character (msec) for each of the four conditions in F&R’s experiment](image)

The long versions of Non-Minimal Attachment sentences took significantly longer to read than long Minimal Attachment sentences, but there was no significant difference between short version of Minimal Attachment and Non-Minimal Attachment. There was a significant difference between long and short versions of Non-Minimal Attachment, but no such difference between long and short versions of Minimal Attachment.

For magnitude estimation, in its on-line web-hosted version adopted here, the prediction is that subjects will estimate (86) and (88) as more complex than (87) and (89), respectively. I expected to find a difference between estimations of long sentences and short sentences. Longer sentences would be estimated as having higher complexity than shorter ones.

2.5.2. Subjects

Sixteen native speakers of English volunteered to participate in the experiment. One subject was excluded from the data analysis; see subsection §2.5.6.2. All subjects were participants of the 2005 LSA summer school at
2.5.3. Materials and design

The sentences of the eye-movement experiment were copied (the full list of sentences is given in Appendix I). There were sixteen items with four versions of each. Examples are provided in (90)-(93) below.

**Minimal Attachment-Long**

(90) I wonder if Tom heard *the latest gossip* about the new neighbors.

**Non-Minimal Attachment-Long**

(91) Tom heard *the latest gossip* about the new neighbors wasn’t true.

**Minimal Attachment-Short**

(92) I wonder if Tom heard *the gossip*.

**Non-Minimal Attachment-Short**

(93) Tom heard *the gossip* wasn’t true.

In the Non-Minimal Attachment conditions, the italicized noun phrase is expected to be erroneously attached as the direct object of the preceding verb, but is found to be the subject of the following verb. This is reflected by having longer reading times. In the Minimal Attachment conditions, the italicized noun phrase is the direct object of the preceding verb, and no slowdown is predicted. The length of the ambiguous noun phrase was varied to control for an effect of ambiguous noun phrase length, yielding a long and a short version of both the Minimal Attachment and Non-Minimal Attachment sentences.

To exclude the possibility that the different versions would differ in the number of clauses, which might lead to higher difficulty in accordance to Minimal Attachment, it was necessary to embed the clause containing the ambiguous phrase in a higher clause in the Minimal Attachment sentences. This manipulation necessitated the use of different lexical items in both versions of the sentences. In other respects, the lexical content of the sentences was matched as much as possible.

The four versions were divided into four lists of sentences. These four lists contained an equal number of items of each version; different versions of a single sentence never appeared within the same list (Latin-Square design). This ensured that each subject read four instances of each of the four versions. In addition to the 16 test sentences, 48 distracter sentences were constructed (not reported in F&R’s study). These sentences had a
variety of different structures balanced in the following way: 12 were long ambiguous sentences known to invoke processing difficulty, 12 long unambiguous sentences, 12 short ambiguous sentences, and 12 short unambiguous sentences. See examples (94) to (97), respectively:

*Long, ambiguous*

(94) The tycoon sold the offshore oil tracts for a lot of money wanted to kill JR.

*Long, unambiguous*

(95) The bank admitted that the mistake would sooner or later cause a lot of trouble.

*Short, ambiguous*

(96) Wherever Alice walks her dog will follow.

*Short, unambiguous*

(97) John figured that students avoid high rents.

The 16 experimental sentences were pseudo-randomized among 48 distracters, so that test items would not appear consecutively.

In order to make sure that subjects comprehended the sentences, a short yes/no question followed each sentence. There were 33 ‘yes’ answers, 31 ‘no’ answers, to avoid possible bias in replies. An example for a question to sentence (86) is *Will the first wife claim the inheritance?* Feedback immediately followed answering. It was not possible to revise an answer once submitted. The subjects had to answer the comprehension question first, and then estimate complexity.

The modulus sentence in all lists was chosen to be a construction known to induce a mild garden path effect (for a discussion, see Starr and Rayner, 2001). The italicized noun phrase in sentence (98) is assumed to be initially processed as a coordinated noun phrase, but it is subsequently realized that the coordination is between clauses, and that the second noun phrase is the subject of the consecutive clause.

*Modulus (NP/CP coordination ambiguity)*

(98) Max was talking to Lisa and the twins were talking to Jane.

Bear in mind that the nature of the modulus is not particularly important nor is its actual numerical estimation, since it functions as a reference point and estimations are relative to it.
2.5.4. Procedure

The web-based questionnaire was divided into three sections: (1) Training section, (2) Practice section, and (3) Experimental section.

In the training section, subjects were asked to estimate the length of a reference line. Five more stimulus lines of varying lengths followed. Subjects had to estimate the lengths of the stimulus lines in proportion to the reference line. For instance, if a subject estimated the reference line to be 4 centimeters, for example, they were supposed to provide the number 12 if the stimulus line seemed 3 times longer than the reference line. The training section is necessary to determine whether, and to demonstrate, that subjects internalized the concept of proportion, the instructions given to them, and could systematically provide estimations of their sensations, rendering their complexity estimations more reliable.

In the subsequent practice section, subjects had to estimate the complexity of a modulus sentence relying on their intuitive judgment. Then they were asked to estimate the complexity of five (or 7, depending on the experiment) more sentences in proportion to the modulus. This section served as practice for the experimental section. It implicitly directed people to understand the concept of proportional sentence complexity judgments, and trained them in manipulations of numerical estimations, so that they would not limit their numerical choices. The third and last section is the experimental one. The task in the experimental section is identical to the one in the practice section.

The instructions given to the subjects are provided in Appendix II. Elaborate instructions were also available for those subjects who wished to read more about the procedure. These are provided in Appendix III.

2.5.5. Apparatus

A web-based Magnitude Estimation experiment was designed using the program WWStim.\(^{27}\) The program allows recording of data over the web by presentation of HTML pages on the subject’s computer. The pages were designed for Microsoft Internet Explorer as it is the most widely spread internet browser. The actual appearance on the subject’s monitor could not be controlled (type of font mostly), but the experiment window was maximized to allow the sentences to appear in one line so that the sentences would not be broken up by the user’s computer. The HTML page could not be resized by the subject.

\(^{27}\) Thanks to Theo Veenker for his help with the program. For specific details about the programming of this technique, see: http://www.let.uu.nl/~Theo.Veenker/personal/wwstim/doc/en.
2.5.6. Results

2.5.6.1. Task comprehension

Individual subjects’ line estimations were divided by the modulus line estimation. The values were then transformed to logarithmic values. The means per each line were regressed with measured logarithmic values of the line proportions.

The regression analysis of the line lengths revealed that $r^2 = 0.99$, the slope: 0.92. The value of $r^2$ means that 99% of the relationship between the log of the measured line length and log magnitude estimations is explained. The slope means that the estimates grew at about 88% as the line length increased.

It can be concluded that the subjects understood the concept of proportion, the instructions given to them, and that they were consistent in their estimations.

2.5.6.2. Reply accuracy

The fifteen subjects answered correctly 81% of the comprehension questions (on average) about the experimental items (Min. = 56%, Max. = 100%). One additional subject was excluded because of a low score of correct answers (43%).

2.5.6.3. Magnitude estimations of complexity

Estimations of each of the test stimuli were standardized by dividing the estimation by the modulus estimation. The standardized proportions were normalized by transforming the standardized values to their logarithmic values. The logarithmic values were averaged over conditions per subject. I will refer to these means as mean complexity estimations. The results are presented in Figure 2 below.
Figure 2: Mean complexity estimations vs. the Minimal Attachment factor for both the long and short versions

Figure 2 shows that the Non-Minimal Attachment condition was always rated to have higher complexity than Minimal Attachment conditions, whether long or short. Also, the long versions were rated to have higher complexity than the short versions.

A 2 (Minimal Attachment versus Non-Minimal Attachment) x 2 (long versus short length) ANOVA was performed on the mean estimations of complexity and indicated that the Non-Minimal Attachment sentences resulted in higher estimations than Minimal Attachment sentences, both in long and short versions. For the Minimal Attachment factor $F_1 (1, 14) = 5.735$, $p = 0.031$, MSe = 0.046; $F_2 (1, 3) = 33.052$, $p = 0.01$, MSe = 0.001; for the length factor: $F_1 (1, 14) = 25.226$, $p = 0.001$, MSe = 0.004; $F_2 (1, 3) = 27.125$, $p = 0.014$, MSe = 0.001. The interaction was not significant: $F_1 (1, 14) = 0.353$, $p = 0.562$, MSe = 0.015; $F_2 (1, 3) = 0.264$, $p = 0.643$, MSe = 0.003.

2.5.7. Discussion

F&R found that the Non-Minimal Attachment condition took longer to read than Minimal Attachment versions, relying on the mean total reading time per character. The interaction between the Minimal Attachment factor and length was significant, and the overall length effect was not significant (this is most likely the result of using reading time per character). No effect of Minimal Attachment was found in the short versions. The reason for that is not discussed. Perhaps this was not F&R’s concern, as they were interested in patterns of eye movement after detection of local structural ambiguity. In
addition, perhaps it is difficult to reliably measure reading times relying on eye movements when manipulated sentences are short. Still, the long version of the Non-Minimal Attachment sentences took longer to read than other versions, relying on the raw numerical values per character.

In the present experiment, Non-Minimal Attachment sentences were estimated significantly more complex than Minimal Attachment sentences, in both the long and short versions. The long versions were estimated as more complex than short versions, and there was no interaction between Minimal Attachment and length.

Both Minimal Attachment and length were significant factors in subjects’ estimation of complexity, but the factors are independent of one another. I therefore conclude that the experiment validates Magnitude Estimation, an offline technique, as a reliable technique for measuring ‘real-time’ effects of Minimal Attachment and length. One advantage of Magnitude Estimation over eye-movements, then, is that it can reliably distinguish between the effect of length and the effect of local structural ambiguity on estimations of complexity.

An alternative interpretation in terms of the thematic parser is also possible. If Non-Minimal Attachment is equated with the effect ‘permitted reanalysis’, then Magnitude Estimation as a technique detects this subtle effect, in addition to the effect of the length of the ambiguous noun phrase.

F&R continued to test for statistical difference of reading times per region. This is impossible in Magnitude Estimation – it does not enable such a measurement. Magnitude Estimation can be easily used as a method for global estimations of complexity with no prediction regarding the locus of complexity. Equally, Magnitude Estimation can be used as a tool for the measurement of complexity without committing to the locus of the detection of difficulty. Nonetheless, one can think of ways of overcoming the confines of Magnitude Estimation: by asking for complexity estimations of regions presented in a moving window paradigm.

To conclude, Magnitude Estimation allows for complex parametric statistical tests, not possible in semantic differential (category scaling), for instance. The technique is accurate and reliably reflects subtle differences in perceived stimuli technically easy and can be conducted over the Internet (A pioneering web-based software called WebExp is offered by Keller et al. for setting up Magnitude Estimation experiments). It remains to be seen what the effect of a web-based experiment on estimations of processing difficulty is in relation to an experiment conducted in a laboratory. Nonetheless, Magnitude Estimation may be used for obtaining quick and reliable results over a large pool of subjects. Since the instructions are easy to comprehend, it can be used to test various populations without the hassle of coming to a laboratory: children, the disabled, the elderly, teenagers, busy
students, working individuals, etc. It can be used also for pre-experimentation in studies of parsing performance and provide a reliable reflection of real-time experimental techniques.

2.6. **Experiment 2: Magnitude Estimation as a tool for the study of processing difficulty in standard garden paths**

2.6.1. **Introduction**

After validating Magnitude Estimation to be adequately sensitive to mild processing effects reflected by ‘permitted reanalysis’, it was required to find out whether it was indicative of processing difficulties which result from a costly local violation of the Theta Principle (as constrained by the OLLC), in accordance with the predictions of the Thematic Transparent Model. The following experiment addresses this question and further attempts to answer the question whether there is a detectable difference (that is, statistically significant difference) in performance between theta-role surplus sentences and other standard garden path constructions, which do not involve theta-role surplus (recall Sections 2.3 and 2.4).

2.6.2. **Subjects**

Twenty native speakers of English volunteered to participate in the experiment. All subjects were recruited over the Internet, and were naïve with respect to the purpose of the experiment. Their average age was 34 years. On average, it took about 30 minutes to complete the experiment (Mean = 28 minutes, St. Deviation = 14). Seven subjects were excluded from the data analysis because they made errors in the line length comparisons. Three more were excluded for not differentiating sufficiently between the sentences (all test estimations were the same).

2.6.3. **Materials and design**

Since there were four conditions, the experiment was divided into two blocks to reduce habituation effects (decrease in sensitivity to the garden path effect upon repeated exposure). One block contained the conditions with object-subject ambiguity (standard garden paths), and their non-ambiguous versions. In the locally ambiguous condition, the initial verb was obligatory transitive; in the disambiguated condition, the initial verb was obligatory intransitive (disambiguation was done by manipulating the thematic lexical entry of the initially encountered verb).

---

28 Thanks are due to [www.language-experiments.org](http://www.language-experiments.org) for publishing the link to the experiment.
Locally ambiguous (standard garden path)

(99) While the policemen were investigating the drug dealers tried to hide the stolen narcotics.

Disambiguated

(100) While the policemen were coming the drug dealers tried to hide the stolen narcotics.

The thematic model predicts that the locally ambiguous condition would result in significantly higher estimations of complexity than its disambiguated equivalent.

The other block contained the theta-role surplus condition and its verb-disambiguated version, the no-surplus condition. In the theta-role surplus condition, the initial verb was obligatorily ditransitive, the second was ambitransitive (either transitive or intransitive). In the disambiguated version, the initial ditransitive verb was replaced by a transitive verb to eliminate theta-role surplus:

Theta-role surplus

(101) The generous host passed the guest-chef who was cooking the smelly and slippery fish.

No surplus

(102) The generous host assisted the guest-chef who was cooking the smelly and slippery fish.

It was predicted that the theta-role surplus condition would be estimated to be more difficult than the no-surplus condition. With regard to the question of interest, whether there is a significant difference between the theta-role surplus and the locally ambiguous condition, it could only be indirectly inferred upon. A direct statistical comparison is impossible, since the two conditions are different in many respects to permit such a statistical test.

The verbs were collected from WebCELEX, which is a web-accessible corpus of written English (and Dutch). Although frequency is not relevant to the experiment, I conducted an analysis of the frequencies of the verbs found in the corpus. The mean frequency (counts per million) of the transitive verbs in the locally ambiguous condition was not significantly different from that of intransitives ($t(4) = 1.332, p = 0.254$, MSe = 382.035). The mean frequency of the ditransitive verbs did not differ significantly from

29 WebCELEX database is run by the Max Planck Institute for Psycholinguistics, see [http://www.mpi.nl/world/celex/](http://www.mpi.nl/world/celex/).
that of the transitives used for the no surplus condition ($t (4) = 1.249, p = 0.280, MSe = 73.843$). The frequency analysis reduces the probability of relating complexity ratings to frequency effects.\(^{30}\)

All sentences were controlled for overall length (about 12 words each). In compliance with the finding of Experiment 1, the length of the ambiguous NP was identical in all conditions. The full list of materials can be found in Appendix IV.

A list consisted of two blocks. The first block contained 6 tokens of the locally ambiguous-disambiguated conditions pseudo-randomized among 18 distracters. The second block contained 6 tokens of the theta-role surplus – no surplus conditions pseudo-randomized among 18 distracters, too (12 test sentences per each questionnaire). This produced 8 lists; the order of blocks was reversed in half of them. The distracters were manipulated in a similar manner as the test sentences, but involved various structures:

**Distractors for the locally ambiguous block**

*Locally Ambiguous*

(103) Until the Jockey settled down the horse that was sleek stood in the big barn.

*Unambiguous*

(104) Sam remembered that our fiftieth wedding anniversary was in November but he missed it.

**Distractors for the theta-role surplus block**

*Locally Ambiguous*

(105) The annoyed doctor warned the senile patient was lying about all that had happened.

*Unambiguous*

(106) I suppose that the clerk knows that the woman wearing that outrageous hat is crazy.

A short yes/no question followed each sentence. There were 12 ‘yes’ answers, 12 ‘no’ answers to avoid bias in replies. The modulus sentence in all lists was a mild garden path given in §2.5.3, repeated here:

---

\(^{30}\) Some verbs not counted in the database were required for disambiguation and for making plausible sentences. The statistical tests reported in the main text therefore exclude the verbs *to observe, to brief, to ignore, to like, to support, to examine, to photograph, to yawn, to come, to idle, to chat, to depart, and to snooze.*
Modulus (NP/CP coordination ambiguity)

(107) Max was talking to Lisa and the twins were talking to Jane.

2.6.4. Procedure and apparatus

The same as specified in §2.5.4 and §2.5.5, respectively.

2.6.5. Results

2.6.5.1. Task comprehension

The regression analysis of the line lengths revealed that $r^2 = 0.99$, the slope: 0.94. It can be concluded that the subjects understood the concept of proportion, the instructions given to them, and that they were systematic in their estimations.

2.6.5.2. Reply accuracy

The twenty subjects answered correctly 88% of the comprehension questions (on average) about the experimental items (Min. = 67%, Max. = 100%). A repeated measures ANOVA on the average correct reply percentages indicated a significant effect of sentence type (locally ambiguous, theta-role surplus, disambiguated, no surplus): $F_1 (2.203, 41.853) = 5.279$, MSe = 673.875, $p = 0.007$. The pairwise comparisons show that the average correct response percentage for the locally ambiguous condition (Mean = 97%, St. Dev. = 10) differed significantly from the theta-role surplus condition (Mean = 72%, St. Dev. = 31), $p = 0.004$. The difference between the disambiguated condition (Mean = 95%, St. Dev. = 12) and the theta-role surplus condition was significant, $p = 0.009$, too. There was no significant difference in the average correct response percentage between the locally ambiguous and disambiguated conditions, and no-surplus conditions (Mean = 87%, St. Dev. = 25), theta-role surplus and no surplus, and disambiguated and no-surplus conditions.

The reply accuracy analysis gave rise to the suspicion that certain types of questions were more difficult than others, which might have affected the results. Presumably, questions that address the ambiguous noun are misunderstood more frequently. The questions were then divided into two groups: a group of questions addressing the ambiguous noun (the drug dealers in (99), for example), and a group of questions that referred to other nouns in the sentences. Seventy percent of all questions were about the

---

31 Since Mauchly’s test of sphericity was significant, the Greenhouse-Geisser corrected F value is reported.
ambiguous noun. Questions not about the ambiguous noun were found only in the control conditions. Since there were no questions that were not about the ambiguous noun in the theta-role surplus condition, it was impossible to determine whether a question that was about an ambiguous noun had an effect on reply accuracy in this condition. For the control conditions, there was no observable reason to assume significant difference in accuracy performance between questions that addressed the ambiguous noun and those that did not (the majority was answered correctly). The conclusion is that questions about the ambiguous noun did not influence reply accuracy, and that another factor must have been acting.

2.6.5.3. Magnitude estimations

The magnitude estimations were calculated in the manner depicted in §2.5.6.3. Since the design consisted of two blocks, only two conditions in each, t-tests were performed (and not an ANOVA). The locally ambiguous condition was significantly different from the disambiguated condition in the subject but not in the item analysis: \( t_1 \) (19) = 2.559, \( p = 0.019\), MSe = 0.034; \( t_2 \) (11) = 1.205, \( p = 0.253\), MSe = 0.049. (The mean for the locally ambiguous condition in the subject analysis was 0.13, for the disambiguated condition 0.04.) A significant difference was found between the theta-role surplus and no-surplus conditions in the subject and item analyses: \( t_1 \) (19) = 6.330, \( p < 0.001\), MSe = 0.026, \( t_2 \) (11) = 2.813, \( p = 0.017\), MSe = 0.058. (The mean for the theta role surplus condition in the subject analysis was 0.19, for the no-surplus condition 0.03).

Given the differences in reply accuracy, items with incorrect replies were discarded, which affected 19% of the data. There was a marginal significant difference between the locally ambiguous and disambiguated conditions in the subject but not in the item analysis: \( t_1 \) (18) = 2.096, \( p = 0.051\), MSe = 0.141; \( t_2 \) (11) = 1.742, \( p = 0.109\), MSe = 0.151. A significant difference was found between the theta-role surplus and no-surplus conditions in the subject and item analyses: \( t_1 \) (15) = 3.012, \( p = 0.009\), MSe = 0.157, \( t_2 \) (10) = 2.312, \( p = 0.043\), MSe = 0.198.

A possible interfering factor in the theta role surplus condition is that some verbs, such as throw, and sell, are very implausible without a final DP attached to them. Others verbs such as, serve, bring, may give up their final DP much more easily (as pointed out by Janet D. Fodor, personal communication). To make sure this is not the explanation for the results, I divided the theta role surplus sentences into two groups. Sentences 2, 3, 5, 6, 9, and 11 (see Appendix IV) with the verbs promise, serve, tell, get, pass, bring (respectively) formed the group of verbs that did not require a final DP. The rest of the sentences were put in a group that required a final DP. In an independent samples t-test, there was no significant difference in
estimations between the two groups of theta role surplus sentences both in the subject and item analyses ($t_1 (10) = 0.158$, $p = 0.878$, MSe = 0.095; $t_2 (10) = 0.847$, $p = 0.417$, MSe = 0.065).  

2.6.5.4. Estimation of occurrence of difficulty

Recall that beside difference in performance between standard garden path constructions and theta-role surplus, Theta Attachment predicts that attachments would be resolved at random. How can random performance be gleaned from the data?

One of the available statistical tests for estimating random coincidence is the $\chi^2$ (chi-square) test. The $\chi^2$ test reports whether the null hypothesis of random occurrence is violated, i.e. whether there was any factor acting other than pure random coincidence. The test requires dichotomous (categorical) data as in the Hebrew questionnaire described in Section 2.4. However, estimations of complexity are continuous (ratio), and not dichotomous. Since only dichotomous (categorical) data is eligible for the $\chi^2$ test, the problem was how to code continuous data as dichotomous data. I did it in two stages.

It was necessary to assess when an occurrence of higher complexity reflected a garden path effect. It was also required to assess what complexity estimations were typical of sentences not supposed to incur processing difficulty.

In order to assess this, I calculated the mean estimations of the disambiguated and no-surplus conditions of each subject. These means served as the baseline of comparison. I compared these means with each item estimation in the locally ambiguous and theta-role surplus conditions, respectively, for each subject. Any estimation in the latter conditions that was higher or equal to the relevant baseline mean was counted as an indication of higher complexity, and thus reflected elevated processing difficulty: therefore, a garden path effect.

Recall that the design was repeated measures, such that there were 3 items for each condition of interest per subject. If more than 50% of the items were counted as difficult, a subject was classified as having experienced higher complexity and therefore difficulty. For example, if a subject indicated higher estimations in 2 out of 3 items in the standard

<table>
<thead>
<tr>
<th>Type of Analysis</th>
<th>Groups of verb types</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Optional final DP</td>
<td>0.50</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Obligatory final DP</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>Item</td>
<td>Optional final DP</td>
<td>0.18</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Obligatory final DP</td>
<td>0.24</td>
<td>0.11</td>
</tr>
</tbody>
</table>

---

$^{32}$ Here are the results of the means for the theta role surplus condition:
garden path condition, then the subject was counted as having experienced higher complexity overall – the subject was likely to have been garden pathed.

The same criterion was applied for the item analysis: if more than 50% of the estimations within the same item of a particular condition were higher or equal to the mean of the baseline, the item was classified as causing higher complexity. For example, if higher estimations were found for more than 50% of the subjects in the standard garden path condition, then the item as a whole was counted as having induced higher complexity. The counts were submitted to a \( \chi^2 \) test of significance.

The locally ambiguous and theta-role surplus conditions were significantly different from random occurrence, i.e. a subject was certain to estimate the test conditions with higher values than the relevant controls. For the locally ambiguous condition, 85% of the subjects were counted as having rated higher values than average, \( p = 0.003 \), in the theta-role surplus condition 95%, \( p < 0.001 \). In the item analysis, 92% of the 12 items in the locally ambiguous condition were counted as being rated with higher values than average, \( p = 0.006 \); and 92% of the 12 items in the theta-role surplus condition were counted to be rated with higher values, too; \( p = 0.006 \).

Another analysis was performed in which estimations of items with incorrect replies were discarded. The locally ambiguous condition was not significantly different from random occurrence: 63% of the subjects were counted as having rated the condition with higher values, \( p = 0.359 \). In the theta-role surplus condition 81% of the subjects were counted as having rated the condition with higher values, \( p = 0.021 \). In the item analysis, 83% of the 12 items in the locally ambiguous condition were counted as being rated higher values, \( p = 0.039 \). Only 67% of the 12 items in the theta-role surplus condition were rated with higher values, \( p = 0.388 \).

2.6.6. Discussion

The questions that were set in the beginning of this experiment were (1) Can Magnitude Estimation detect processing difficulty related to garden path constructions as defined within the Thematic Transparent Model? (2) Can a difference in performance between theta-role surplus sentences and standard garden path sentences be detected?

The significant differences in estimations across conditions in the subject analysis indicate that Magnitude Estimation as a method can detect differences in complexity, which is equal to processing difficulty (these terms denote the same thing in garden path constructions). The item analysis of the locally ambiguous and disambiguated conditions shows that not all test items induced difficulty in equal measure. This is not the case in the theta-role surplus and no-surplus conditions (all items induced equal
measures of difficulty). Nonetheless, it was possible to detect reliable differences in estimations between the manipulated conditions over subjects.

The theta-role surplus condition had a significantly lesser degree of reply accuracy than the other conditions (only 72% correct). The result indicates subjects’ awareness of the attachment ambiguity of the final PP, which might have lead to an increase in estimations of this particular condition. Just from performance on reply accuracy, one may conclude that there is a difference in performance between theta-role surplus and locally ambiguous conditions (97% correct).

Why were subjects prone to estimate the theta-role surplus condition with higher values? As Theta Attachment does not resolve local ambiguity in such cases, the theta-role surplus condition was estimated with higher values, as it was difficult to accurately answer the comprehension questions. This was not the case in the locally ambiguous condition, where subjects answered with high accuracy and estimated it with lower values. This is because attachment ambiguity is easier to resolve as Theta Attachment does adjudicate. These results imply that participants modulated their judgments of difficulty with the questions.

However, considering the results of the item analysis, the picture becomes clearer. The locally ambiguous conditions always induced higher ratings, presumably reflecting difficulty, whereas the theta-role surplus condition induced difficulty at random.

2.7. Experiment 3: Theta-role surplus in Dutch

2.7.1. Introduction to part A of the experiment

Experiment 2 showed a trend compatible with the prediction that performance on the theta-role surplus condition differs from performance on standard garden paths. A more direct statistical comparison is required, where both types of sentences (theta-role surplus and standard garden paths) will only differ in a manipulated factor, therefore permitting submission to a reliable and direct statistical test. Dutch allows such a comparison, and the following experiment applies it to the question at hand.

Part A of experiment 3 was exploratory and was dedicated to finding out whether processing difficulties were robust enough in Dutch to be detected by Magnitude Estimation. Part B of the experiment allowed concluding that performance on theta-role surplus condition was different from performance on the standard garden path condition and the unproblematic condition. Further analysis suggests that performance is at random on theta-role surplus conditions.
2.7.2. Subjects
Forty native speakers of Dutch participated in the experiment for course credit. All subjects were naïve with respect to the purpose of the experiment. Their average age was 21.38 (95%) were speakers of standard Dutch. 2 (5%) spoke standard Dutch and next to it had command of a regional dialect. On average, it took about 20 minutes to complete the experiment (Mean = 22 minutes, St. Deviation = 10). Six participants were excluded from the statistical analysis for making mistakes in the line comparison.

2.7.3. Materials and design
In the theta-role surplus condition, e.g. sentence (108) below, the post verbal PP ‘over de looproutes’ about the trail can receive the <THEME> theta-role from the ditransitive ‘vertellen’ tell or the <PATIENT> role from ‘gelezen’ read.

(108) Theta-role surplus
De begeleider vertelde de wandelaarster die al
The guide (Mas) told the hiker (Fem) who already
heel wat had gelezen over de looproutes.
quite a lot had read about the trail.
‘The guide told the hiker who had already read quite a lot about the trail.’

Sentence (109) is an example of the locally ambiguous condition.

(109) Locally ambiguous
De begeleider vertelde de wandelaarster die met haar
The guide (Mas) told the hiker (Fem) who with her
schoenen had geruild over de looproutes.
shoes had swapped about the trail.
‘The guide told the hiker who had swapped shoes with her about the trail.’

The critical point is ‘haar’ her, which is locally ambiguous between a possessive form and a pronoun. Under Theta Attachment, ‘met’ with (a preposition with a thematic role) will initially give its theta-role to ‘haar schoenen’ her shoes. Upon reaching the verb ‘geruild’ swapped, there is no argument to pMerge with the thematic position of the verb. A costly local violation of the Theta Principle arises which results in processing difficulty. Subsequent uncostly restructuring requires the PP ‘met haar schoenen’ with her shoes to become ‘[pp met haar] [np schoenen]’ to obtain a grammatical
The predictions of all other models are the same as summarized in Section 2.3.

The experiment design requires a control condition, termed Baseline A: the locally ambiguous condition disambiguated by a determiner, see example (110). The insertion of a determiner between 'met haar' with her and 'schoenen' shoes indicates that only 'haar' her is pMerged with the preposition head, since *'met haar de schoenen' with her the shoes is not a grammatical option. Difficulty is not predicted to occur in this condition.

(110) Baseline A
De begeleider vertelde de wandelaarster die met haar de schoenen had geruild over de looproutes.
The guide (Mas) told the hiker (Fem) who with her the shoes had swapped about the trail.

'The guide told the hiker who had swapped the shoes with her about the trail.'

Both conditions involve the same lexical noun phrases. The linear sequence of phrases in (108) is DP VP DP Comp AdvP VP PP. The sequence in (109) is: DP VP DP Comp PP/(PP NP) VP PP. An AdvP is not predicted to cause any difficulty and its length was counterbalanced by the PP in (109).

In (108), the ambiguity is attachment of the PP to the relevant licensing verbs; in (109) between attachment of a PP or a PP NP. The sentences differ only in the ambiguity of attachment of the prepositional phrase, all other things being equal.

The overall length of the sentences was controlled for (approximately 14 words per sentence). In compliance with the finding of Experiment 1, the length of the ambiguous PP was controlled for, and was identical in all conditions (P DP NP). The length of the PP-dominated NP in the locally ambiguous condition was also controlled for and was always P Pos/NP NP. All sentences contained the same number of clauses. The full list of sentences is given in Appendix V.

The three conditions were divided into three lists of sentences. Each list contained an equal number of items of each condition; different conditions of a single item never appeared within the same list (Latin-Square design). The 12 test sentences were pseudo-randomized among 36 distracter sentences, which were controlled for overall length. These had a variety of different structures divided in the following way: 12 were sentences known

---

33 Note that the difficulty of (109) cannot be associated with the possibility of having several potential antecedents for 'haar' her, since the only feminine antecedent that matches the feminine feature of the pronoun is the NP that c-commands the relative clause. All other antecedent nouns are masculine.
to induce difficulty, e.g. (111); 12 were sentences that did not induce difficulty (112); and 12 were disambiguated garden path constructions (113):

**Examples of distracters**

(111) *Locally ambiguous*

De speelgoed soldaat smolt langzaam en gelijkmatig als was hij van chocolade droop hij weg.

'Locally ambiguous: The toy-soldier melted slowly and evenly as if it were made of chocolate.'

(112) *No ambiguity*

De wat oudere beeldhouwster werkte verrassend genoeg aan een standbeeld van haar overleden echtgenoot.

'No ambiguity: The somewhat elderly sculptress worked surprisingly enough on a statue of her deceased husband.'

(113) *Disambiguated*

Gek genoeg wedijverde de zwemkampioene die getrouwd was met haar coach met een schaatster.

'Disambiguated: Oddly enough the swimming champion who was married to her coach competed with a skater.'

A short yes/no question followed each sentence to test comprehension. There were 18 ‘yes’ answers and 18 ‘no’ answers, to avoid bias in replies. An example for a question to sentence (108) is ‘Kreeg de wandelaarster uitleg over de te lopen routes?’ *Did the hiker get an explanation about the trail?* Feedback immediately followed answering. The subjects had to answer the comprehension question first, and then estimate complexity. It was not possible to revise an answer once submitted.

The modulus sentence in all lists was a theta-role surplus sentence, which in a pilot test induced a mild processing difficulty. Note that the nature of the modulus is not particularly important nor is its actual numerical estimation, since it functions as a reference point and estimations are made in proportion to it.
(114) *Modulus*

Nog steeds verlangde de vrouw die was vertrokken
Still longed the woman who had left
naar een schuilplaats naar haar vaderland.
for a hiding-place for her country.

‘The woman who had left for a hiding-place still longed for her country.’

2.7.4. Plausibility and acceptability pre-tests

I conducted a plausibility pre-test to ascertain that the post verbal PP in the theta-role surplus condition could be plausibly attached to both the (high) ditransitive and (low) ambitransitive verb. It was also necessary to see whether the verbs in each condition possessed the thematic properties associated with them (transitivity, intransitivity, ambitransitivity) in an acceptability experiment.

The conclusion from the plausibility and acceptability pre-tests was that the final PPs in the theta-role surplus sentences could be plausibly attached to both the lower and upper verbs. The lower verb could omit the final PP, but retained plausibility (viz. the lower verb is ambitransitive). The thematic properties of the verbs as classified were confirmed, too. Other possible attachments were excluded to eliminate the possible complication of considering other attachment possibilities than intended. This ensured that considerations of plausibility and acceptability of attachments would not bias attachment decisions. For an elaboration on the plausibility and acceptability experiments and their statistical results, refer to Appendix VI (the instructions for the plausibility and acceptability test are found in Appendix VII).

2.7.5. Procedure and apparatus

The procedure and apparatus were the same as described in §2.5.4 and §2.5.5. The instructions in Dutch are given in Appendix VIII, and the detailed instructions in Appendix IX.

2.7.6. Results

2.7.6.1. Task comprehension

The regression analysis of the line lengths revealed that $r^2 = 0.99$, the slope: 0.92. It can be concluded that the subjects understood the concept of proportion, the instructions given to them, and that they were systematic in their estimations.
2.7.6.2. Reply accuracy
The 40 subjects correctly answered 92% of the questions (on average) about the experimental items (Min. = 83%, Max. = 100%). There was no significant difference between the average correct response percentages across conditions (all Fs < 1). This implies that the questions did not interact with estimations of complexity.

2.7.6.3. Magnitude estimations of complexity
The magnitude estimations were calculated in the same manner as described in §2.5.6.3. In a repeated measures ANOVA of sentence type (locally ambiguous, theta-role surplus, Baseline A), both the subject and item analyses were significant: $F_1 (2, 78) = 9.469, p < 0.001$, $MSe = 0.009$; $F_2 (2, 22) = 7.960, p = 0.003$, $MSe = 0.005$.

The pairwise comparisons in the subject analysis showed a significant difference between the locally ambiguous and theta-role surplus conditions, and between the theta-role surplus condition and Baseline A, $p < 0.001$. The baseline condition was rated higher than the locally ambiguous condition: the mean of the locally ambiguous condition in the subject analysis was 0.03, the theta-role surplus: -0.05, the Baseline A: 0.02. There was no significant difference between the locally ambiguous condition and Baseline A, $p = 0.580$.

The pairwise comparisons in the item analysis patterned similarly. A significant difference between the locally ambiguous and theta-role surplus conditions, $p = 0.012$; and between the theta-role surplus condition and Baseline A, $p = 0.005$. There was no significant difference between the locally ambiguous condition and Baseline A, $p = 0.369$.

2.7.7. Discussion
The point at hand is that there was a significant difference in mean complexity estimations between locally ambiguous and theta-role surplus conditions. The difference in difficulty can only be attributed to the manipulated factor of ambiguity of attachment. At any rate, a garden path effect in Dutch is measurable by Magnitude Estimation.

An unexpected result of this experiment is that baseline A was estimated not to be different from the locally ambiguous condition; the mean of the latter was numerically lower than that of the Baseline A. The reason for this could be that determiners, being orthographically short, go unnoticed during reading. Subjects simply did not notice them, and were primed to equate the disambiguated versions with the ambiguous ones, thus providing higher estimations. The next experiment was set about to correct the problem of disambiguation.
2.7.8. Introduction to part B

The same set of sentences as in part A of the experiment (see §2.7.1) was used, excluding baseline A. Another method for disambiguation is the inclusion of commas, which people naturally indicate as necessary in cases of garden path constructions. However, disambiguation by a comma of the locally ambiguous condition would have been unnatural:

(115) Unnatural disambiguation by a comma

De begeleider vertelde de wandelaarster die met haar,
The guide (Mas) told the hiker (Fem) who with her,
schoenen had geruild over de looproutes.
shoes had swapped about the trail.

??'The guide told the hiker who had swapped shoes with her, about the trail.'

To avoid possible complications of unnaturalness, a comma disambiguated version of the theta-role surplus condition was used. The positions of the commas was natural (see example (118) below), since they demarcated a clause and thus disambiguated the sentences.

2.7.9. Subjects

Thirty-seven native speakers of Dutch participated in the experiment for course credit. All subjects were naïve with respect to the purpose of the experiment. Their average age was 20, 34 (92%) were speakers of standard Dutch, 3 (8%) standard Dutch and fluent in a regional dialect. On average, it took 20 minutes to complete the experiment (Mean: 22 minutes, St. Deviation: 8). The results of two participants were removed because of making mistakes in the line comparison and another two for lack of differentiation between test sentences.

2.7.10. Materials and design

The theta-role surplus and locally ambiguous conditions were the same as reported in §2.7.3. Baseline B consisted of the theta-role surplus condition disambiguated by commas. The sentences were tested for plausibility and acceptability as laid out in §2.7.4. The full list of sentence is found in Appendix X.
(116) *Theta-role surplus*

De begeleider vertelde de wandelaarster die al heel wat had gelezen over de looproutes.

‘The guide told the hiker who already read quite a lot about the trail.’

(117) *Locally ambiguous*

De begeleider vertelde de wandelaarster die met haar schoenen had geruild over de looproutes.

‘The guide told the hiker who had swapped shoes with her about the trail.’

(118) *Baseline B*

De begeleider vertelde de wandelaarster, die al heel wat had gelezen, over de looproutes.

‘The guide told the hiker, who had already read quite a lot, about the trail.’

Note that in order to avoid priming of the commas added to Baseline B, a comma was added to a third of the distracter sentences, too.

2.7.11. **Procedure and apparatus**

See §2.7.5.

2.7.12. **Results**

2.7.12.1. *Task comprehension*

The regression analysis of the line lengths revealed that $r^2 = 0.99$, the slope: 0.91. It can be concluded that the subjects understood the concept of proportion, the instructions given to them, and that they were systematic in their estimations.

2.7.12.2. *Reply accuracy*

The 37 subjects correctly answered 90% of the questions (on average) about the experimental items (St. Dev. 8, Min.: 67%, Max.: 100%). There was no significant difference between the average correct response percentages across conditions ($F (2, 72) = 1.449, p = 0.242, MSe = 283.721$).
Magnitude estimations were calculated in the same manner mentioned in §2.5.6.3. In a repeated measures ANOVA of sentence type (locally ambiguous, theta-role surplus, Baseline B), both the subject and item analyses were significant: $F_1(1.691, 61.064) = 17.815, p < 0.001, \text{MSe} = 0.021$; $F_2(2, 22) = 11.665, p < 0.001, \text{MSe} = 0.009$.

The pairwise comparisons in the subject analysis showed a significant difference between the locally ambiguous and theta-role surplus conditions ($p = 0.006$), between the theta-role surplus condition and Baseline B, $p < 0.001$; and between the locally ambiguous condition and Baseline B, $p < 0.001$. This is summarized in Figure 3.

Since Mauchly’s test of sphericity was significant, the Greenhouse-Geisser corrected $F$ value is reported.

Figure 3: Mean magnitude estimations of complexity per condition

The pairwise comparisons in the item analysis showed a similar pattern. There was a significant difference between the locally ambiguous and theta-role surplus conditions, $p = 0.046$; between the theta-role surplus condition and Baseline B, $p = 0.011$; and between the locally ambiguous condition and Baseline B, $p = 0.001$. 

---

34 Since Mauchly’s test of sphericity was significant, the Greenhouse-Geisser corrected $F$ value is reported.
2.7.12.4. Estimations of occurrence of difficulty

The occurrence of higher values was calculated in the manner described in §2.6.5.4. The occurrence of higher values in the locally ambiguous condition was significantly different from random coincidence: 73% of the 37 subjects rated the condition with higher values, \( p = 0.008 \). In the theta-role surplus condition, 57% of the subjects rated this condition higher, \( p = 0.511 \), which means random occurrence of higher values.

In the item analysis, 92% of the 12 locally ambiguous items resulted in higher values, which was significantly different from random coincidence, \( p = 0.006 \). In the theta-role surplus items, 67% of the 12 items invoked higher values, \( p = 0.388 \); i.e. an item of the theta-role surplus condition resulted in higher values at random.

2.7.13. Discussion

Part B of experiment 3 shows that there is a significant difference in performance between the theta-role surplus and the locally ambiguous conditions. All subjects fully comprehended the sentences in the various conditions as shown by the high number of correct replies. Therefore, the difference in performance was not reflected in reply accuracy as in Experiment 2. It was reflected by magnitude estimations. Furthermore, the comma disambiguation method did not go unnoticed. As Figure 3 demonstrates, Baseline B was estimated the lowest. The locally ambiguous sentences were estimated to be most complex, and the theta-role surplus condition was estimated in between the most difficult locally ambiguous condition and the easiest, disambiguated condition.

This result is compatible with the predictions of thematic parsing only and not with any other processing model. The random occurrence of higher values in the theta-role surplus condition is compatible with the thematic parser’s prediction that attachments are resolved at random. Subjects were consistent in providing higher values in the locally ambiguous condition, and lower values in the Baseline B condition. Although this outcome was estimated from the data obtained, it nonetheless provides further confirmation to the predictions of the thematic parser, and particularly to Theta Attachment.

2.8. Conclusion: Theta Attachment as a guiding principle of human parsing performance

This chapter was concerned with the question whether Theta Attachment is a guiding principle of parsing. Furthermore, it wished to ground the Thematic Transparent Model with empirical evidence and strengthen its ‘goodness of fit’ or, in other words, its conformity to external reality reflected by
processing difficulty. A theoretical analysis yielded a specific prediction of the model, which set it apart from other relevant models. The specific prediction was that in certain cases, local attachment ambiguity is resolved at random. These are cases of theta-role surplus, which allow two possible alternatives – both equally satisfy the Theta Principle. Choosing one alternative leads to processing difficulty, the other does not. The distribution of occurrence of difficulty across subjects and items would be random in these cases. The answer to this question could only be inferred from the data given methodological limitations (see §2.6.5.4) and was formulated statistically as whether there was any factor other than random coincidence which might have affected the occurrence of processing difficulty, which was taken to be reflected by higher values of magnitude estimations.

Another experimental prediction was a significant difference in performance between theta-role surplus and standard garden path and non-garden path constructions.

I chose to use the Magnitude Estimation methodology for acceptability judgments, given that garden path constructions are locally ungrammatical. Experiment 1 successfully validated Magnitude Estimation, an off-line technique, as a methodology that could detect mild processing difficulty ('permitted reanalysis') much like real-time experimental techniques. Magnitude Estimation was reliable in the disentanglement of length effects and mild processing difficulties, and even yielded better results than tracking of eye-movements.

Experiment 2 was concerned with performance on theta-role surplus constructions and standard constructions in English that contain local structural ambiguity. When Theta Attachment did not resolve local ambiguity as in the case of theta-role surplus, this condition was estimated most difficult, though reply accuracy scores were much lower than those in the canonical garden path condition. In the latter case, where Theta Attachment does resolve local ambiguity, subjects estimated tokens as easy, a conclusion warranted by the higher accuracy in reply scores.

With regard to the question which prediction was borne out, experiment 2 can only be regarded as inconclusive, since the reply accuracy performance complied with the thematic model’s main prediction, whereas the estimation of occurrence of difficulty seemed to contradict it. At any rate, the experiment gave rise to the reasonable suspicion that the predictions of thematic parser were on the right track. Since a direct statistical comparison between the two conditions was impossible, a clearer and more direct comparison was desired.

Experiment 3 looked for a direct comparison of the conditions of interest in Dutch. After overcoming experimental problems in Part A of the experiment, Part B established that canonical garden path (locally
ambiguous) sentences were estimated most complex, i.e. most difficult, and their comma disambiguated sentences as easiest. The mean of the theta-role surplus condition was exactly in the middle between the means of the two other conditions. Furthermore, it was indicated that the theta-role surplus was estimated difficult (=higher) at random. The disambiguated version was estimated easy, and the locally ambiguous condition most difficult.

These results are compatible only with the predictions of the Thematic Transparent Model and Theta Attachment, both in terms of occurrence of difficulty and means of estimations (recall the two types of predictions in Section 2.3). By that, they provide empirical support to the claim that the human sentence processing mechanism is guided by Theta Attachment.

The theoretical analysis of constructions involving local structural ambiguity and processing difficulty leads to the conclusion that they demonstrate local ungrammaticality – they locally violate the Theta Principle. Inversely, the inductive analysis of processing difficulty corresponds with predictions of the deductive analysis. It is logical then that to the extent that phenomena of processing difficulty conform to a certain model, then the model itself reflects this external reality, which is reflected by a measurable processing difficulty in the brain (as in ERP, e.g. Osterhout, L. McLaughlin, J. Bersick, M. (1997)). The model satisfies the requirement of ‘goodness of fit’. Insofar as processing difficulty conforms to local violation of the Theta Principle, it can be said that the principle is real, at least of the phenomena at hand, and is therefore represented in the human mind.
3. CENTER-EMBEDDINGS

3.1. The importance of center-embeddings

In computational linguistics, computational models are considered successful if they show to have satisfactory ‘goodness of fit’, i.e. they explain the corpus of data they were supposed to explain. A model’s ‘goodness of fit’ is further improved if it can be shown that the principles it incorporates are also generalizable to a set of data it was originally not designed to account for. Thus, there are two measures for choosing the best model, its ‘goodness of fit’, and its ‘generalizability’ (Pitt and Navarro, 2005).

Within the field of parsing performance, one of the most prominent phenomena of processing difficulties (in addition to garden paths) is center-embeddings. A natural set of data for testing, then, the generalizability of the Thematic Transparent Model is center-embeddings.

My purpose in this chapter is to show that the principles of the thematic parser, established on the basis of the reasonable ‘goodness of fit’ between it and processing difficulties of garden path constructions, are generalizable to center-embeddings. Put differently, the Thematic Transparent Model as proposed accounts for processing difficulty associated with center-embeddings, too. As an upshot, the account further establishes

55 They use generalizability as the ability to replicate the model on various different sets of data.
the psychological reality of Theta Attachment and the Theta Principle.

Generalizability is also justified on grounds of ‘parsing adequacy’. A model of human parsing performance should account for all or most of requirements (119)-(123) or at least demonstrate the feasibility of such account within the model (Vosse and Kempen, 2000).

(119) Complexity of linguistic input (center-embedding vs. left/right branching).
(120) Syntactic ambiguity (global and local attachment ambiguities).
(121) Lexical ambiguity (multiple word senses).
(122) Lexical and syntactic priming (facilitation due to recent use of a lexical item or a syntactic structure).
(123) Semantic context effects (thematic fit, subcategorization effects, pragmatic fit, plausibility).

Factors (120)-(123) were handled in the previous chapters. For a more elaborate discussion of lexical ambiguity, refer to Pritchett (1992a), Section 4.1. It should be noted that lexical ambiguities are regarded there as syntactic phenomena. This chapter concentrates on (119).

I will discuss differences in processing between left-branching and right-branching languages as well. There is a persistent myth in linguistic circles that speakers of left-branching languages can process one more level of center-embedding than speakers of right-branching languages (as confirmed by Jean Aitchison, personal communication). Apparently, the capacity of short-term memory is bigger if one happens to speak a left-branching language. The myth is traceable to Rosenbaum and Kim (1976), but will be demonstrated to be based on an error in comparison. The difficulty of center-embeddings in right-branching languages cannot be compared to an identical left-branching structure of canonical center-embeddings. There is no difference in the capacity of short-term memory.

### 3.2. Center-embeddings as islands: a new observation

The classical definition of center-embeddings traces back to the olden days of transformational grammar. The phrases A and B form a nested construction (or center-embedding) if A falls totally within B, with some non-null element to its left within B and some non-null element to its right within B (Chomsky, 1965). Thus the phrase the man who wrote the book that you told me about is nested in the phrase called the man who wrote the book that you told me about up in I called the man who wrote the book that you told me about up. Center-embeddings are also referred to as nested constructions.
The phrase A is said to be self-embedded in B if A is nested (center-embedded) in B and, furthermore, A is a phrase of the same type as B. Thus who the students recognized is self-embedded in who the boy who the students recognized pointed out in the man who the boy who the students recognized pointed out is a friend of mine, since both are relative clauses. Thus, center-embedding has to do with bracketing, and self-embedding with labeling of brackets as well. Every self-embedding is center-embedding, but not every center-embedding is self-embedding.

To appreciate the difficulty that center-embeddings induce, consider first the following sentences:

(124) The mouse died.
(125) The mouse (that) [the cat bit] died.

The relative clause (that) the cat bit was center-embedded between the mouse and died in (124). The resulting sentence (125) is grammatical and easily processed. However, center-embedding (equally, self-embedding) the relative clause (that) the dog chased in (125) results in severe processing difficulty:

(126) The mouse (that) the cat (that) [the dog chased] bit died.

Relying on (126), Miller and Chomsky (1963) concluded that the accumulation of grammatical operations, self-embedding in this case, leads to processing difficulty. The operations manifested in it belong to grammatical knowledge and therefore grammar does not exclude it as ungrammatical (now in the regular sense):

“Note that it would be quite impossible to characterize the unacceptable sentences in grammatical terms. For example, we cannot formulate particular rules of the grammar in such a way as to exclude them.” (Chomsky, 1965, p. 11)

Rather, the high number of self-embeddings leads to processing difficulty: the capacity of human beings’ short-term memory is exceeded when there are more than two self-embeddings. The observation was supported by an experiment that showed that humans could not process more than two self-embeddings of the type given here (Marks, 1968). So important and convincing were these findings (although Miller and Chomsky’s influential work has not been assessed (Tomalin, 2006)), that self-embedding of object relative clauses became a canonical example of center-embeddings and the

---

36 It has been claimed that the rarity of center-embeddings is the source of difficulty. This has been shown to be misguided (for a review, see De Roeck et al., 1982).
term ‘center-embedding’ became synonymous with such constructions. The construction served as rare and important evidence for the capacity of the human mind in processing sentences:

“...We must conclude that the competence of the native speaker cannot be characterized by a finite automaton. The grammar stored in his brain cannot be a one-sided linear grammar, a fact that is not in the least surprising. Nevertheless, the performance of the speaker or hearer must be representable by a finite automaton of some sort. The speaker-hearer has only a finite memory, a part of which he uses to store the rules of his grammar (a set of rules for a device with unbounded memory), and a part of which he uses for computation in actually producing a sentence or “perceiving” its structure and understanding it.” (Chomsky, 1963, p. 390)

Concomitant to Miller & Chomsky’s observations, center-embeddings have been a focus of psycholinguistic research, but not so much of syntactic-theoretical study. It was accepted that the difficulty of center-embeddings is a processing phenomenon, and therefore they were not to be studied within grammar. Relating to the quote above, the task of psycholinguistics was to provide an account why center-embeddings take a finite automaton out of its processing abilities. Over the years since, psycholinguistic hypotheses suggested various processing factors in center-embeddings, which might drive short-term memory to its limit. One of the prominent contemporary theories, for example, is Similarity-Based Theory (Lewis, 1996). It argues that the initial NPs deficient of syntactic markings may all potentially assume the same syntactic position such that they interfere with one other – increasing difficulty.

Along with the above-mentioned goals, the purpose of this chapter is to show that the difficulty associated with center-embeddings can be predicted by the local application of global grammatical principles during parsing. Two conceptual barriers need to be overcome. The first is that center-embeddings are amenable to an account that excludes them on syntactic grounds closely related to the manner in which parsing is carried out.

The second change relates to the nature of the parser. As argued in the quote above, performance was describable by a finite automaton of some sort. Canonical center-embeddings demonstrated that performance was a finite automaton, because the accumulation of center-embeddings used up processing capabilities.

In the continuation, I will argue that processing difficulty of center-embeddings is a local violation of the Theta Principle, ultimately a failure to
Center-embeddings comprehend the sentence. Based on this argument, I claim that canonical center-embeddings cannot serve as evidence for limitations of short-term memory. The present explanation does not seek to exclude short-term memory as a necessity for the processing of linguistic input or deny its limited capacity, but rather asks whether it is necessary to invoke it at all as a direct cause for the difficulty of center-embeddings, given the explanatory consistency of the Thematic Transparent Model.

In brief, I argue that short-term memory (describable by a finite automaton) is not domain-specific, but domain general. It is a prerequisite for computations of linguistic input or any other type of information, but it is insufficient for explaining the processing difficulty in center-embeddings.

The argument rests on a theory-independent observation, unnoticed until now and is a discovery of some sort, concerning the syntactic status of such constructions. Center-embeddings can be regarded as a complex NP island embedded in a complex NP island (as in Ross, 1967):

(127) The [Complex NP island mouse the [Complex NP island cat the dog chased] bit] died.

Note that this observation cannot be dismissed or contended upon simply because it is a statement of fact. Once this observation is made, it fundamentally changes the way center-embeddings should be investigated. Center-embeddings can now be related to the study of grammar, much like any other island. The immediate question that springs to mind is whether processing difficulty associated with these constructions is due to locality considerations defined within grammatical theory and suggested for other instances of islands, such as the complex NP island. As will be demonstrated, the reply is an unequivocal yes – perhaps surprisingly so because it goes against what has been thought about these constructions in the past fifty years.

To provide experimental grounds to the observation that the difficulty of center-embeddings might be related to the status of center-embeddings as complex NP islands, Experiment 4 was constructed. It provides experimental evidence that factors considered early on in the generative tradition to affect performance on center-embeddings are not at the core of explanation of the difficulty of center-embeddings nor are they sufficient.

### 3.3. Experiment 4: The origin of difficulty of center-embeddings

#### 3.3.1. Introduction

Experiment 4 seeks to find out the source of difficulty of center-embeddings out of a collection of hypotheses that were made about these constructions
early on in generative grammar. It aims to compare various approaches to the origin of difficulty of center-embeddings suggested over the years.

3.3.1.1. Predictions of Rosenbaum and Kim (1976)

Rosenbaum and Kim (1976) made the point that gradation of difficulty of self-embedded constructions correlated with the number of relativizations within a given construction and with the number of embeddings, both appeared to consume roughly an equal amount of resources. Consider the following examples, which their structure they discuss (ORC = Object Relative Clause):

(129) The warden forwarded the demand [that the prisoner [ORC that the social worker visited] made].
(130) The report [that the prisoners [that the social worker visited] ran away] was alarming to many citizens.
(131) The disclosure [that the report [that the prisoners ran away] was true] alarmed many citizens.
(132) The naughty boy threw the stone [that hit the window [that the neighbor fixed the day before]].

Sentence (128) contains two self-embeddings (of relative clauses) and two object relativizations, and is the canonical center-embedding example. Sentence (129) has one self-embedding, and two object relativizations. Sentence (130) incorporates two self-embeddings and one object relativization, while sentence (131) has only two self-embeddings, but no object relativizations. The phrases in sentence (132) are sequential; there is no self-embedding or relativization. Rosenbaum and Kim predicted that the gradation of difficulty would be (128) > (129) = (130) > (131) > (132), and confirmed this gradation in a provisory grammaticality judgment experiment.

The gradation of difficulty was not tested in a controlled experiment that could be accurately replicated. The additional purpose of the experiment reported here is to test whether this gradation of difficulty is indeed borne out under replicable conditions.

37 Relativization was considered a transformation at the time, consuming more cognitive resources.
3.3.1.2. Predictions of Miller and Chomsky (1963)

If one strictly adheres to Miller and Chomsky, i.e. that accumulation of self-embeddings is the only factor which influences difficulty, then this leads to the following order of difficulty: (128) = (130) = (131) > (129) > (132). This is due to the nature of self-embeddings: too many dependencies need to be kept in short-term memory until completed. The observation was later expressed by Kimball (1973), given also the experimental evidence of the time:

(133) *The principle of two sentences:* The constituents of no more that two sentences can be parsed at one time.

3.3.1.3. Predictions within the Thematic Transparent Model

Given the observation of islands, the view here is that constructions involving a complex NP island embedded in another complex NP island would correlate with the highest difficulty. Since a complex NP island inevitably contains a relative clause, this property can also be regarded as a relative clause embedded in a relative clause. Sentences (128) and (129) incorporate such property, and would therefore be the most difficult to process. Sentences (130) and (131) will be easier, because they do not incorporate such property. Sentence (132) will be easiest. In summary, the prediction of difficulty is (128) = (129) > (130) = (131) = (132).

The goal of experiment 4 was two-fold then: (a) to test whether the number of self-embeddings and/or object relativizations had a significant effect on perceived difficulty; (b) to check which gradation of difficulty was borne out under replicable conditions.

3.3.2. Subjects

Forty subjects volunteered to participate in a Magnitude Estimation experiment over the Internet. Their average age was 36.

3.3.3. Materials and design

Examples of the test conditions are provided in (134)-(138). Special care was given to the use of the same nouns in all items of the various conditions. Where this turned out to be impossible, each token within a condition remained in the same semantic field in the other condition. The overall average number of words per sentence was 14; the sentences maintained equal overall length.
Two self-embeddings and two object relativizations (2SE2OR)

(134) The defendant that the lawyer that the district-attorney appointed cross-examined incriminated himself in the trial.

One self-embeddings and two object relativizations (1SE2OR)

(135) The defendant consulted the lawyer that the district attorney that the judge warned appointed.

Two self-embeddings and one object relativization (2SE1OR)

(136) The fact that the defendant that the district-attorney cross-examined incriminated himself was worrying.

Two self-embeddings (2SE)

(137) The fact that the report that the defendant incriminated himself was true was posted everywhere.

Sequential

(138) The district attorney appointed the lawyer that cross-examined the defendant that incriminated himself several times during the trial.

The full list of test items is found in Appendix XI. Two items per condition were distributed among five lists in a Latin-Square design. The test items were pseudo-randomized among 36 fillers so that no two test items occurred one after the other. The fillers were matched with the test items in overall length and complexity. An example of a filler sentence is found in (139).

Filler

(139) The daughter of the king’s son hurt himself in a competition and was rushed off to the royal hospital.

The modulus sentence for all lists was sequential:

Modulus

(140) The dog chased the cat that bit the mouse that desperately tried to run away to safety.

3.3.4. Procedure and apparatus

The procedure and apparatus were the same as specified in §2.5.4 and §2.5.5. The subjects were instructed to rate the ease of comprehension; the instructions are found in Appendix XII.
3.3.5. Results

3.3.5.1. Task comprehension

The results of the regression analysis for the line length comparisons were $r^2 = 0.99$, the slope was 0.88. The regression analysis shows that the participants internalized the concept of proportion and understood the instructions, rendering their sentence magnitude estimations reliable.

3.3.5.2. Magnitude estimations

A repeated measures ANOVA for sentence type (2SE2OR, 1SE2OR, 2SE, 2SE1OR, Sequential) was conducted on the standardized mean estimations.

Figure 4: Standardized mean estimations per condition

Significance was found for the sentence type factor: $F_1 (4, 156) = 5.349, p < 0.001$, MSe = 0.229, $F_2 (4, 36) = 2.991, p < 0.031$, MSe = 0.104. The results of the pairwise comparisons are summarized in Table 1.
### Table 1: Pairwise comparisons

<table>
<thead>
<tr>
<th>Pair</th>
<th>Std. Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SE2OR - 1SE2OR</td>
<td>0.063</td>
<td>0.275</td>
</tr>
<tr>
<td>2SE2OR - 2SE1OR</td>
<td>0.092</td>
<td>0.011</td>
</tr>
<tr>
<td>2SE2OR - 2SE</td>
<td>0.127</td>
<td>0.014</td>
</tr>
<tr>
<td>2SE2OR - Sequential</td>
<td>0.138</td>
<td>0.004</td>
</tr>
<tr>
<td>1SE2OR - 2SE1OR</td>
<td>0.178</td>
<td>0.036</td>
</tr>
<tr>
<td>1SE2OR - 2SE</td>
<td>0.117</td>
<td>0.034</td>
</tr>
<tr>
<td>1SE2OR - Sequential</td>
<td>0.127</td>
<td>0.009</td>
</tr>
<tr>
<td>2SE1OR - 2SE</td>
<td>0.111</td>
<td>0.482</td>
</tr>
<tr>
<td>2SE1OR - Sequential</td>
<td>0.102</td>
<td>0.100</td>
</tr>
<tr>
<td>2SE - Sequential</td>
<td>0.087</td>
<td>0.292</td>
</tr>
</tbody>
</table>

With regard to the self-embedding factor, it is concluded from the difference between 2SE2OR and 1SE2OR that it did not have a significant effect ($p > 0.05$). Since the difference between 2SE1OR and 2SE was not significant, it is concluded that the factor of object relativization did not significantly affect estimations. The difference between 2SE2OR and 2SE1OR was significant, but this cannot be due to the number of object relative clauses (there was no significant difference between 2SE1OR and 2SE), it can only be related to the fact that a complex NP island is embedded in a complex NP island as found in 2SE2OR. Otherwise, the difference between 2SE and 2SE1OR should have been significant. At any rate, there is a clear cut between the group of conditions that contain two object relativizations and conditions that lack two object relativizations. The difference among the conditions with two object relativizations was not significant, and the difference among the rest of the conditions was not significant, too. Only the difference between conditions with two object relativizations and those without was significant.

From the pairwise comparisons, and Figure 4, where 2SE2OR and 1SE2OR were estimated highest than the other sentence types, it can be concluded that the order of difficulty is $(134) = (135) > (136) = (137) = (138)$ in compliance with the predictions of the Thematic Transparent Model.

#### 3.3.6 Discussion

There is good evidence that when object relative clauses are embedded in one another (essentially a complex NP island is embedded in another complex NP island), estimations of difficulty increase in relation to other self-embedded structures.

As for the claim that accumulation of self-embeddings is the main cause of difficulty, no evidence was found for that. Note that this does not
mean by itself that the 2SE condition is as easy as the Sequential condition –
their respective means were different – only that this does not amount to a
significant difference. That is, there may be some effect of number of self-
embeddings, but this fails to explain the difference between 2SE2OR and
2SE1OR.

Questions remain concerning other factors suggested in
psycholinguistic literature, which might affect difficulties. Experiment 5 in
Section 3.4 examined predictions of two other theories particularly designed
to account for the difficulty of center-embeddings.

3.4. Experiment 5: The Similarity-Based and Dependency Locality
Theories

3.4.1. Introduction to the various predictions

3.4.1.1. Predictions of Similarity-Based Interference Theory

The theory of Similarity-Based Interference (Lewis, 1996, 1998, Lewis and
Nakayama, 2001) argues that the difficulty associated with center-embedded
structures is the amount of similarity of interfering factors between two
dependents. Difficulty is measured by the combined effects of retroactive
and proactive interference on syntactic attachments. Consider the sequence
in (141), where V is a verb that must retrieve its NP argument:

(141) φ₁ φ₂ φ₃ ... φₙ NP ρ₁ ρ₂ ρ₃ ... ρₙ V

The association of V with NP suffers from ‘retroactive interference’ from the
intervening items ρ₁ ρ₂ ρ₃ ... ρₙ and ‘proactive interference’ from the prior
items φ₁ φ₂ φ₃ ... φₙ that are still active in the parse. The theory asserts that
the critical factor determining the amount of interference is the similarity of
the interfering items to the to-be-retrieved (to-be-attached) item NP.
Similarity is taken to be syntactic; in essence, the syntactic role the NP may
assume. For example, if the NP is to be assigned the structural position of a
subject, then there is proactive interference from those items in φ₁...φₙ that
have not yet been attached and may fill a subject position. In addition, there
is retroactive interference from those items in ρ₁ ... ρₙ that were potential
subjects. Consider sentence (128), for instance, repeated here:

(142) The reporter who the senator who John met attacked disliked the
editor.

Upon processing the verb attacked, the NP John causes retroactive
interference because it assumed a subject role as the intended NP the
senator. Why does John cause interference if it has already been attached to
meet at the time attacked got into the input? It is assumed that the interfering
noun causes ‘storage interference’. Additionally, the NP the reporter causes proactive interference, because it may still assume a subject role. In short, proactive interference results from retrieval difficulties, and retroactive interference from storage interference. Thus, canonical center-embeddings manifest high interference.

Consider next the pair (143) and (144) below. Sentence (143) has more interference than sentence (144). The pair differs only in the syntactic elements of the most embedded clause.

(143) The report that the suggestion [that the fact was accurate] was misleading was contradictory.

(144) The report that the suggestion [that the landlord evaded taxes] was misleading was contradictory.

Nouns such as report, suggestion, and fact typically select clausal complements. Therefore, the NP the landlord causes less retroactive interference when was misleading is encountered in (144), whereas in (143) the fact causes retroactive interference upon encountering was misleading. Furthermore, the initial noun phrases the report and the suggestion proactively interfere as potential subjects when was accurate is encountered in (143). In (144), however, at the verb evaded, there is no proactive interference of the initial NPs (the report or the suggestion are less likely to evade, being inanimate).

This set of sentences was interesting to test in an experiment to see whether the factors suggested by the Similarity-Based Interference Theory influenced difficulty. Sentence (143) has more interference than (144), and this might be reflected by higher difficulty.

These sentences were combined with another group of sentences that tested the predictions of the Dependency Locality Theory, to which I turn next.

3.4.1.2. Predictions of the Dependency Locality Theory

The Dependency Locality Theory (DLT) (Gibson, 2000) assumes that the difficulty of center-embeddings reflects complexity of discourse. It suggests a metric for estimating this complexity, based on two principles. The first is the ‘discourse processing cost’, which captures the cost associated with accessing or constructing the discourse structure for the maximal projection of the input head $h_2$ (a noun or verb heads):

(145) **Discourse processing cost:** 1 Energy Unit (EU) is consumed if $h_2$ is the head of a new discourse referent; 0 EUs otherwise.

The second principle captures the structural integration cost associated with connecting the syntactic structure for a newly input head $h_2$ to a projection of
a head \( h_1 \) that is part of the current structure for the input. This cost is dependent on the complexity of the computations that took place between \( h_1 \) and \( h_2 \).

(146) **Structural integration cost:** 1 EU is consumed for each new discourse referent in the intervening region.

For example, the canonical center-embedding is predicted to be most difficult at the verb *bit*:

(147) The mouse that the cat that [the dog chased] bit died.

The discourse processing cost of *bit* itself is 1 EU, *bit* being a head of a new discourse referent. The structural integration of *bit* with its subject *the cat* costs 2 EUs, since two discourse referents intervene: *the dog*, and *chased*. Furthermore, the direct object of *bit* must form a dependency with *that*, the complementizer. Since the discourse referents *bit*, *chased*, *the dog*, and *the cat* intervene between the trace and *that*, the structural integration cost is 4 EUs. In total, then, the complexity of integrating *bit* is 7 EUs, a high measure of complexity. The DLT predicts that (147) will be difficult to process due to this high measure of complexity.

Let us consider the processing of the following paired examples:

(148) The claim that the rumor [that the guard snoozed] spread reached the headlines.

(149) The claim that the rumor [that the guard was lazy] spread reached the headlines.

As far as the calculation of EUs, the relevant difference between the pair of sentences resides in the most embedded clause. In (148), there will be 1EU for the discourse processing cost of *snoozed* as it is an eventive verb, and 1EU for its structural integration with *the guard* (two extra EUs in total). In (149), no EUs will be associated with the discourse processing cost of *was lazy* – there is no event, it is a stage. This item is not a discourse referent (‘discourse referent’ is defined as an entity that has a spatiotemporal location so that it can later be referred to with an anaphoric expression, such as pronouns for NPs, or tense on a verb for events). The prediction is that (148) will be more complex than (149), all other things being equal.

3.4.1.3. Predictions of the Thematic Transparent Model

Sentences (143)-(144) and (148)-(149) do not have factors that might ameliorate or exacerbate the difficulty associated with center-embeddings, since they all have the same syntactic status. It is therefore the prediction of the thematic parser that no substantial difference in difficulty will be found among those conditions (the null-hypothesis).
3.4.2. Subjects
Thirty subjects volunteered to participate in the experiment over the Internet. Their average age was 26 years old.

3.4.3. Materials and design
Examples of the sentences used for testing the predictions of the Similarity-Based Theory are given in (150) and (151). Sentences such as (150) were titled ‘high interference;’ sentences of the type in (151) were titled ‘low interference’.

High Interference
(150) The announcement that the rumor [that the claim was shocking] was true was exciting.

Low Interference
(151) The announcement that the rumor [that the actress won an award] was true was exciting.

The sentences for testing the predictions of the DLT are provided in (152)-(154). Prior to testing, it was noted that sentences such as (152) (Event Short Version) were shorter in length than (154) (Stage). To factor out length, another condition was added: the Event Long Version. This version was identical to Event Short Version, but an adverb was inserted in the most embedded clause. The insertion of an adverb is not predicted to incur difficulty by the DLT since it is not a discourse element. Thus the prediction was that ESV and ELV would be more complex than Stage.

Event Short Version (ESV)
(152) The claim that the rumor that the mayor quit circulated surprised voters.

Event Long Version (ELV)
(153) The claim that the rumor that the mayor unexpectedly quit circulated surprised voters.

Stage
(154) The claim that the rumor that the mayor was a gangster circulated surprised voters.

Three items per condition were distributed into five lists in a Latin-Square design, and were pseudo-randomized among 48 fillers so that no two test items occurred one after the other. The fillers from Experiment 4 were matched with the test items in overall length and complexity. The full list of
experimental items is found in Appendix XIII. The modulus was sequential, the same one as in experiment 4 (for examples of the modulus and fillers, refer to §3.3.3).

3.4.4. Procedure and apparatus

The procedure and apparatus are the same as in experiment 4 (see § 3.3.4). The experiment was published in “The portal for psychological experiments in language” (http://www.language-experiments.org).

3.4.5. Results

3.4.5.1. Task comprehension

The results of the regression analysis are $r^2 = 0.99$, the slope is 0.95. The participants were therefore reliable in their concept of proportion and understood the instructions.

3.4.5.2. Magnitude estimations

All of the results rely on the log-transformed standardized mean estimations, calculated in the manner explained in §2.5.6.3.

There was no significant difference between the conditions of High and Low Interference: $t_1 (29) = 1.028, p = 0.313; t_2 (14) = 1.351, p = 0.198$.

A repeated measures ANOVA of sentence type (ESV, ELV, and Stage) was conducted. No significant overall effect was found: $F_1 (2, 58) = 1.059, p = 0.354, MSe = 12.763; F_2 (2, 28) = 0.987, p = 0.385, MSe = 0.005$.

3.4.6. Discussion

No evidence was found that length, discourse structure, or syntactic interference manipulated in the most embedded clause had an impact on estimations of difficulty. Although in Experiment 1 (Section 2.5) a significant effect was found for length, apparently a difference of one word does not amount to significance. It means that all sentences were difficult but not because of the length or the manipulated factors. Perhaps the difficulty of the structure self-embeddings masks the difficulty of length (in compliance with Experiment 4, where there was an indication that the number of self-embeddings might have a significant effect).

That no evidence was found is not to mean that Magnitude Estimation as a method is not sensitive enough to these factors. As has been demonstrated in Chapter 2, there is good evidence that Magnitude Estimation is sensitive to length, at least. The conclusion above is therefore correct: there is no reason to suspect that the factors that Similarity-Based Interference Theory and the DLT offer have an influence on estimations.
The experiment is not a very strong argument in favor of the thematic parser, of course, because it only entails the acceptance of the null hypothesis of this particular model (no difference among conditions). Still, this null hypothesis cannot be rejected, which complies with the Thematic Transparent Model.

The findings are insufficient as a basis for rejecting the Similarity-Based Interference Theory and the DLT as possible account of the source of processing difficulty in center-embeddings. Further evidence is required. Section 3.5 sets to explore the principles of the DLT, and Section 3.6 of the Similarity-Based Interference Theory.

3.5. Notes on the Dependency Locality Theory

A legitimate question concerning the DLT would be whether the concept of structural integration in (146) gives the right results and whether it captures the facts known about locality.

The structural integration cost is approximated by the number of discourse referents between a new head and its dependent, which has already been integrated. However, in example (147), the element the dog can be considered to be the subject of bit, as this integration is more local than the dependency between the cat and bit, and even felicitous (dogs bite). Further, in sentence (155) taken from Gibson (2000), higher complexity is at the verb hope: 4 EUs (1 EU discourse integration, plus 3 EUs for the intervening discourse referents: the editor, sent, and the photographer).

(155) The reporter who the photographer sent to the editor hoped for a good story.

However, an additional possible linear structural integration can be made between hope and the photographer, yielding the reduced relative ambiguity typical of garden paths the photographer (who was) sent to the editor hoped for a good story.

It is not clear why the grammatically correct dependency is chosen, as is evident from the examples provided in Gibson’s paper, and more local felicitous dependencies are not. Plainly, no constraint is provided on permissible dependencies.

Far more disconcerting is the fact that intervening material within a dependency need not be problematic. There are known examples in which intervening discourse referents do not give rise to elevated complexity. Long distance dependencies are not complex in relation to center-embeddings, although many discourse referents interfere in this dependency.18

18 It cannot be argued that the DLT is only valid for nested constructions, and therefore does not cover the
(156) Who did Felix say that John heard that Lisa whispered that Mary liked?

The question in (156) can be answered quite accurately and easily (namely, that Mary liked, say, the guy next door). However, according to the structural integration cost, associating the trace of like with who would consume 8 EUs, corresponding to the intervening discourse referents liked, Mary, whispered, Lisa, heard, John, Felix, and say. Question (156) is predicted to be more complex than (147), which is obviously not the case. The structural integration cost makes the wrong prediction here.

Considering the fact that the DLT was primarily designed to account for the complexity of nested constructions, it can be further inquired whether it has reasonable ‘goodness of fit’, i.e. whether it accounts for the corpus of data of nested constructions. As Gibson (1991) himself reports, there are nested constructions that do not induce difficulty:

(157) What the woman that John married likes is chocolate.

The complexity at likes amounts to 7 EUs (1 EU for the discourse integration of like, 2 EUs for John and married, which intervene between like and the woman, and 4 EUs between the trace of like and what, corresponding to 4 discourse referents: likes, married, John, and the woman). It is therefore erroneously predicted that the complexity of (157) would be equal to the canonical (147). However, Gibson (2000) offers no explanation how example (157) would fit the DLT or, alternatively, why the DLT might not be applicable to it.

Furthermore, consider the highly difficult center-embedding example (158):

(158) Chickens chickens chickens peck peck peck.

Meaning: chickens which peck (other) chickens, which peck other chickens too, (tend to) peck.

The sentence is a nested construction, short, and still very difficult to understand. The structural integration cost on the final peck depends on four discourse referents only. The DLT erroneously predicts low degree of difficulty (compare also to the easy (160)): the complexity on the third peck is 5 EUs (1 EU for discourse integration cost, and 4 EUs for structural integration, corresponding to the intervening chickens and peck), but the sentence is in fact very difficult to comprehend.

Complexity of long distance dependencies. As noted, one of the tests for a good theory is its generalizability. If a theory cannot be generalized to a set of data which has a direct bearing on the principles it suggests, although it was not designed for it, then the theory is at fault.
Given examples (157) and (158), the DLT does not have ‘goodness of fit’, even for constructions it was meant to explain. In the former case, it predicts difficulty where there is none; in the latter, it predicts no difficulty where it does occur.

In fact, both principles suggested in the DLT, the discourse integration cost and the structural integration cost, are stipulative – an independent motivation is not offered why complexity, approximated by the number of discourse referents intervening in a linear fashion between two dependents, is a proper measure of complexity, and why integrating them into discourse is complex. The author claims that it “turned out” to be so, but offers no compelling argument for this matter.

An apparent justification for suggesting the DLT is presented as the following. Sentence (160) is ameliorated by substituting John in (159) with first person pronoun.

(159) The reporter who the senator who John met attacked disliked the editor

(160) The reporter who the senator who I met attacked disliked the editor.

In the beginning of Gibson’s paper, readers are told that it has been reported in literature (in the now famous Bever (1970) paper) that the substitution of a proper noun (or DP) with 1st person or 2nd person pronouns results in amelioration. A complexity rating experiment is reported in the DLT paper, which was conducted to confirm what was previously observed and to indicate that none of the previous theories about complexity explains this type of amelioration. This observation provided some impetus for suggesting the DLT. Since 1st and 2nd pronouns are not discourse referents, they are not considered in the calculation of the structural integration cost. At the point of attacked in (160), the integration cost is only 5 EUs, compared to 7 EUs in (159) – fewer discourse referents are crossed. The explanation thus “provides a straightforward account of the observed contrast” (Gibson, 2000, p. 107) and the DLT is suggested as superior to other theories. However, it would be surprising if the theory did not explain this. The DLT took amelioration by pronouns as its merit. It does not explain this type of amelioration, the theory simply presupposes it.

Moreover, there appears to be a contradiction in the explanation. On one hand, it is suggested that “many resource complexity effects can be explained using integration cost alone” (ibid, p. 102). On the other hand, “the comprehension difficulty at a word in a sentence is assumed to be

---

39 Recall that ‘discourse referent’ is defined as an entity that has a spatiotemporal location so that it can later be referred to with an anaphoric expression, such as pronouns for NPs, or tense on a verb for events.
determined” by (1) frequency factors; (2) structural integration; (3) the storage cost; (4) contextual plausibility; (5) the discourse complexity; (6) reanalysis difficulty; and (7) overall intuitive complexity (comprehension difficulty at a word in a sentence is integration cost, as the discussion in the original paper reveals). Thus, the pair of principles in (145) and (146) are an “oversimplification” (ibid., p. 105) of the overall complexity of a certain word. It is not clear then how one should approximate complexity: by the pair of principles suggested, or by all of the other factors. Each word must be considered on a single case basis, such that the approximation of complexity is so complex in itself that generalization is not possible, although everything is explainable – it’s simply complex.

3.6. Notes on the Similarity-Based Interference Theory

Although the Similarity-Based Interference Theory may perhaps capture a grain of truth concerning the elevated difficulty readers encounter when faced with a sequence of identical NPs (in terms of their syntactic role), it suffers from the same defects of the DLT.

The theory makes the wrong predictions with regard to embedded sentential subjects (157) (and embedded free relatives (177) – a discussion follows), erroneously predicting difficulty where there is none. In the case of long-distance dependencies (156), it predicts an explosion of retroactive and proactive interference (when *likes* must attach its object), that the sentence will be incomprehensible, contra fact.

In addition, consider the following Japanese example (from Vermeulen, 2005).

(161) *Multiple possessive construction in Japanese*

Kitahankyuu-ga anettai-ga usagi-ga mimi-ga naga-i.

North Hemisphere-Nom subtropics- Nom rabbit-Nom ear-Nom long-Present

‘It is the North Hemisphere, where rabbits in the subtropics have long ears.’

The expectation is very high degree of difficulty when *naga-i* is encountered, since the parser has to determine the structural role of the elements marked by –*ga* (these could be possessors, Nominative markers, etc.), which might lead to a high degree proactive interference. However, the sentence does not incur any known difficulty (Vermeulen, personal communication) and these types of sentences did no receive any attention from psycholinguistics as far as I know.

It should be emphasized that the phonological repetition of –*ga* markers in Japanese or their functional ambiguity are not the single source of difficulty of center-embeddings (Uehara, 2003). Also, the repetitive nominative marking is not the source of difficulty (Vasishth, 2001). The cost
associated with 3ga sequences appears only when the surface-adjacent constituents are in a double center-embedding construction: “3ga cost is specific to double embedding” (Uehara, 2003, p. 335). Sequences of 3ga cannot be the whole story. At any rate, the Similarity-Based Interference Theory simply does not reflect the core reason of difficulty of center-embeddings.

After excluding factors suggested in literature as possible origins for the difficulty of center-embeddings, an alternative account within the Thematic Transparent Model is suggested in the remainder of this chapter. The following sections explore the origin of difficulty of various center-embedded constructions reported in literature, in further satisfaction of requirement (119). It will be argued that the difficulty of center-embedding is related to grammar and involves local violations of the Theta Principle.

3.7. Center-embedded object relative clauses & the thematic parser

Consider again an example of canonical center-embeddings: center-embedded object relative clauses as in (162):

(162) The mouse the cat the dog chased bit died.

Next, I follow the parsing steps as guided by Theta Attachment.

(Step a) **The mouse, the cat, the dog:** the three initial constituents are identified as determiner phrases, and are stored in short-term memory. The following table organizes the input constituents, the elements in ‘store’ (unattached elements) and the syntactic structure built in store. The table makes no claim on ‘psychological reality’ and is used for expository reasons only.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Input constituents</th>
<th>Store (unattached elements)</th>
<th>Resulting structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>the mouse</td>
<td>the mouse</td>
<td>DP the mouse</td>
</tr>
<tr>
<td>2.</td>
<td>the cat</td>
<td>the mouse, the cat</td>
<td>DP the cat</td>
</tr>
<tr>
<td>3.</td>
<td>the dog</td>
<td>the mouse, the cat, the dog</td>
<td>DP the dog</td>
</tr>
</tbody>
</table>

(Step b) **chased:** the element *chased* is identified as an obligatory transitive verb and its maximal theta grid is retrieved from the lexicon, containing two theta-roles (external and internal). The DP *the dog* is identified as an
argument and is pMerged with the subject [Spec, TP] position in local satisfaction of Theta Attachment.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Input constituents</th>
<th>Store (unattached elements)</th>
<th>Resulting structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>the mouse</td>
<td>the mouse</td>
<td>DP the mouse</td>
</tr>
<tr>
<td>2.</td>
<td>the cat</td>
<td>the mouse, the cat</td>
<td>DP the cat</td>
</tr>
<tr>
<td>3.</td>
<td>the dog</td>
<td>the mouse, the cat, the dog</td>
<td>DP the dog</td>
</tr>
<tr>
<td>4.</td>
<td>chased: &lt;θ₁, θ₂&gt;</td>
<td>the mouse, the cat, θ₁</td>
<td>TP chased</td>
</tr>
</tbody>
</table>

(Step c) **bit**: the verb *bit* is introduced, and is identified as an obligatorily transitive verb. Its maximal theta grid is retrieved, containing two theta-roles (external and internal). According to Adjacency, it becomes clear now that a noun phrase would not appear in the direct object position of *chased*. Forced by Theta Attachment, the empty direct object position of *chase* is licensed by the internal theta-role. This licensed thematic position must Merge with an argument to satisfy the Theta Principle. Recall that the Theta Principle requires Pure Merge (Merge without Move). In order to satisfy the requirement for Pure Merge, an operator must be merged with the [Spec, CP] position (recall that a full CP is projected when a finite verb is encountered, see §1.5.1.2). In the licensed thematic position, the Theta

---

Note that several indications conspire to show that there is indeed a relative clause. First, there are two consecutive NPs (an impossibility in English unless the sequence is a reduced relative clause). Second, an argument NP is missing within the VP of the relative CP, as indicated by Theta Attachment and required by Adjacency. Third, the head noun of the relative clause is present and requires being associated with a theta-role. If a complementizer is overt, it might serve as stronger indication that the C-head is a head of a relative clause. All of these indications follow from the assumption that the parser is transparent, and has grammar installed.
Principle is satisfied when an operator binds it. The operator ranges over the head noun (*the cat*) and restricts its interpretation.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Input constituents</th>
<th>Store (unattached elements)</th>
<th>Resulting structure</th>
</tr>
</thead>
</table>

5. bit: <0₁, 0₂> the mouse, 0₁, 0₂ (both of bit)

It is important to understand the distinction between thematic *licensing* and thematic *binding* (Higginbotham, 1985). An argument is thematically licensed, when it is pmerged with a thematic position (which is in turn licensed by a theta-role that is part of the lexical entry of a predicate). Thematic licensing is a property of the base position. Head nouns of (restrictive) relative clauses are thematically bound, but not thematically licensed (or simply, head nouns may not pMerge).

Bear in mind that (restrictive) relative clauses c-commanded by head nouns incorporate the requirement that their thematic operator bind a thematic position (receiving the values of the θ-features of the thematic base position), in satisfaction of the Theta Principle. Thematic binding is established by the Merge of a thematic operator with [Spec, CP]. The thematic operator binds the values of the θ-features (as within the Theta System (Reinhart, 2002), i.e., the values of [c] and [m], see §1.5.1.4), which are licensed by the theta role of the verb at the thematic position within the relative clause. The operator associates these thematic values with the head noun. I shall term C-heads possessing the requirement for thematic binding as relativizing heads.
In the parsing steps that follow, I will adopt the Probe-Goal Framework (Chomsky, 1999, 2000). The thematic operator misses its values upon Merge and functions as a Probe, which searches the closest Goal in its domain (its c-command domain). The Goal is the values of the θ-features at the licensed thematic position (for ease of exposition, I will sometimes refer to the Goal as simply a theta role). Both Probe and Goal are active in the sense that the former requires to be valued, and the latter needs to be bound.

The head of CP is a relativizing head. Like the thematic operator, it functions as a Probe, which searches in its domain the closest (values of the θ-features of the) thematic position, the Goal. For independent syntactic reasons within the Probe-Goal framework, a Goal must be in an edge position (Specifier) to enable binding by a Probe (Chomsky, 2005). The relativizing head attracts the values of the θ-features to the Specifier position of the VP to allow binding by the thematic operator. ‘Thematic’ Attract operation is reserved for object relative clauses only (see elaborations in Section 3.15 and §3.10.2). Now the thematic operator can bind the closest thematic values.41 (Note that these operations occur immediately to locally satisfy Theta Attachment as soon as possible. Thus, the operations described in this step are still part of stage 5 in the table.)

...
Now that the relative NP *(the) cat the dog chased* is thematically interpretable, it is merged with the external thematic position of *bit*. Since no new input is introduced at this stage, this operation still belongs to stage 5 in the following table.
(Step d) **died**: the element *died* is entered to the parse. Forced by Theta Attachment and Adjacency, the remaining theta-role of *bit* licenses its direct object position, in the same manner described in stage (c), and a thematic operator is merged with [Spec, CP].
As before, the thematic relativizing C\textsubscript{2}-head attracts (the values of the \( \theta \)-features of) \( \theta_k \) to the edge. \( Op_k \) must now bind the closest theta-role. The Goal would have been \( \theta'_k \) (a copy of the values). However, \( \theta'_k \) is not found in its search domain. Because \( Op_i \) c-commands \( \theta'_k \), \( Op_i \) is closer to \( \theta'_k \) than \( Op_k \). The former operator blocks binding by the latter, as demonstrated by the temporary parse representation in Tree 14 below (the blocking operator is marked by a double-lined box). This situation is similar to an intervention effect (Relativized Minimality (Rizzi, 1990) or Minimal Link Condition (Chomsky, 1995)), where a local structural relation cannot hold between a Probe (\( Op_i \)) and a Goal (\( \theta'_k \)), if an element (\( Op_k \)) which is a potential bearer of the relevant relation intervenes between the Probe and the Goal.
This surprising situation relies on the following definition of c-command:

(163) **C-command**: $\alpha$ c-commands $\beta$ if (a) every $\gamma$ that dominates $\alpha$ dominates $\beta$; and (b) $\alpha$ and $\beta$ are disconnected. ($\alpha$ and $\beta$ are disconnected if $\alpha \neq \beta$, and neither dominates the other. *Domination* is taken to be an irreflexive relation where $\alpha$ dominates $\beta$ if every segment of $\alpha$ dominates $\beta$ (Chomsky, 1995, Kayne, 1994)).

C-command between $Op$ and $\theta'_i$ is established in the following way. [Spec, $C_iP$] is disconnected from [Spec, VP], as neither dominates the other. The DP *the* and the NP *cat* are both two segment categories which do not dominate $Op$, for domination is a requirement on every segment of a phrase (not every segment of the DP or the NP dominates $Op$). $CP_1$ does not dominate $Op$, since domination is irreflexive. $Op$ is dominated by $CP_2$; the latter also dominates $\theta'_i$. ($Op$ is as if it were located in the Specifier position.
of TP, the subject position.) Consequently, $Op_i$ c-commands $\theta'_i$ ($\theta'_i$ is dominated by the TP so it cannot c-command $Op_i$). 42

The upper thematic operator $Op_i$ has no (values of $\theta$-features of) theta role to bind, the consequence of which is that the NP mouse has no thematic association within the clause. Further, the licensed thematic position of bit is left unbound (essentially, a theta role has no argument to Merge with). These are local violations of the Theta Principle, which is impossible to satisfy under the given circumstances of intervention. Consequently, the sentence is extremely difficult to comprehend. (Note that the inability of the parser to pMerge the external theta-role of die with an argument NP is not considered problematic, given that there is no grammatical possibility to locally satisfy the Theta Principle: the relative NP the mouse... was not formed.)

The intervention of the operator of the most embedded relative NP with the satisfaction of the Theta Principle is the key finding of the discussion of the parsing steps. The local inability to satisfy the Theta Principle because of intervention is the core reason for humans’ experience of difficulty of canonical center-embeddings. It is a consequence of the structure the parser has formed. Surprisingly, grammar can exclude such structures. It is this discovery that offers an explanation to “those [center-embeddings, O. S. L] most mystical of linguistic beasts” (Schütze, 1996, p. 57).

The parsing of (162) was guided by the intuition that difficulty arises due to interference between a thematic Probe and a thematic Goal; where another closer, more local, thematic element causes interference. This interference does not allow for the satisfaction of the Theta Principle, yielding processing difficulty, which is also difficulty of comprehension. The situation described above does not exactly conform to the traditional configuration of intervention. Such structure is demonstrated in Tree 15 below, where $Op_j$ and $Op_i$ are thematic Probes and $\theta_j$ is a thematic Goal.

42 The relative NP is not headed by a DP, as is normally assumed in current syntactic theory (Thanks are due to Josef Bayer for bringing up this question). This structural choice is justified on two grounds. First, it provides an explanatory gain and leads to a natural account of the contrast between canonical center-embeddings and ameliorated constructions (see in the continuation in the main text). Crucially, because of the undeniable observation that canonical center-embeddings incorporate islands, this structural choice enables an explanation which relies on an intervention effect as is currently assumed for islands in competence theory.

Furthermore, as Section 3.9 argues, ‘binding out’ is a property of being in a Specifier. Since the most embedded complex NP island is in a Specifier, it binds out.
The Probe $Op_i$ blocks $Op_j$ from binding the intended $\theta_j$, given that it is a potential bearer of the thematic relation between $Op_i$ and $\theta_j$. The ‘uninterpreted’ or ‘valued’ $\theta$-features in the base position $\theta_j$ cannot be copied and interpreted in the [Spec, CP] position. Furthermore, $\theta_j$ is licensed but not bound: the role is not associated with an argument. The Theta Principle is locally dissatisfied and processing difficulty ensues.

The finding of intervention simply relies on the manner with which natural language processing is executed: incrementally (or, paratactically), from left-to-right. Intervention is a consequence of this manner of computation, and is intricately related to grammatical relations formed during parsing.

Noam Chomsky (personal communication) points out that the mere island status of center-embeddings is insufficient for explaining the difficulty of (162) as might appear to have been claimed in Experiment 4. Sentence (164) contains embedded complex NP islands, but is relatively not difficult to process.

(164) The [complex NP man who thinks that the [complex NP cat that the dog chased yesterday] bit the mouse] is very angry.

What is required to experience processing difficulty in center-embeddings is intervention that disables the local satisfaction of the Theta Principle, and not only their island status. Experiment 4 only claimed that the higher difficulty in canonical center-embeddings is associated with islands, but did not make the claim that difficulty is due to islands. The observation that center-embeddings incorporate islands only allowed for asking whether intervention in terms of locality considerations (Relativized Minimality (Rizzi, 1990) and/or the Minimal Link Condition (Chomsky, 1995)) does occur.
We may now explain why intervention is not possible in (164). First, the thematic Probe $O_{pi}$ is dominated by the CP that is projected by the sentential complementizer that, and therefore may not c-command out.\footnote{Since the structural position of yesterday is somewhat controversial, Tree 16 omits its projection. This has no consequence to the explanation.}

Second, as shown in Tree 16, there is no potential Goal for the Probe $O_{pi}$. As demonstrated in (Step c) above, the parser, constrained by Adjacency, may

---

**Tree 16**

Since the structural position of yesterday is somewhat controversial, Tree 16 omits its projection. This has no consequence to the explanation.
only license a thematic position when it is clear that no argument NP will appear, i.e. when it is clear what the post-verbal constituent is. However, in (164), the post-verbal constituent of *bit* is an overt argument NP: *the mouse*, which readily undergoes pMerge. The Probe $Op_i$ does not have a Goal in the shape of a thematic role it can bind and interfere with binding by $Op_j$.

### 3.8. Subject relative clause embedded in an object relative clause

The intervention effect found in canonical center-embedded object relative clauses occurs also when a subject relative clause, a complex NP island too, is embedded in an object relative clause.

(165) ?The fisherman that the killer-whale that chased the seal bit died.

The parsing steps are conducted in the same manner of (164).

(a) **The fisherman, that, the killer-whale, that**: Abiding to Theta Attachment, the initial constituents *the fisherman, the killer-whale* are identified and project as determiner phrases, and the instances of *that* are identified as complementizers. Since no thematic information relevant for attachments has yet appeared, the elements are stored.

(b) **chased**: The element *chased* is identified as an obligatorily transitive and its maximal theta grid is retrieved, containing two theta-roles (external and internal). The [Spec, TP] position is licensed by the external thematic role of *chased* (recall example (28) in §1.5.1.2). Next, a thematic operator is merged with the [Spec, CP] position to allow binding and creating a subject relative NP.

(166) The fisherman that [DP the [NP killer-whale [CP1 $Op_i$ that $\theta_i$ chased]]].

Note that since the external theta-role has already been attached to an edge position, an Attract operation by the local thematic relativizing C$_1$-head is not required. The thematic operator binds the co-indexed values of the features of the theta-role. Having the relative NP roleless does not carry any special load on the parser because it is grammatically impossible to allow for the satisfaction of the Theta Principle.

(c) **the seal**: This constituent is introduced and is identified as an argument. The maximal theta grid of *chased* can be locally satisfied now, purely merging *the seal* with the internal theta position. The situation is as demonstrated by the following representation:

(167) The fisherman that [DP the [NP killer-whale [CP1 $Op_i$ that $\theta_i$ chased the seal]]].

(d) **bit**: The verb *bit* is identified as an obligatorily transitive verb and its maximal theta grid is retrieved having two theta-roles. The relativized DP
the killer-whale that chased the seal can be merged with the external thematic position of bit.

(e) **died**: Forced by Theta Attachment and Adjacency, the internal theta-role of bit may now license its direct object position.

The thematic relativizing C₂-head attracts θk to the edge. And now Op₂ must bind the values of the features of its co-indexed theta-role. Observing Tree 17 below at this crucial point of the parse reveals the same intervention effect as in (162).

Tree 17

```
      DP
     /  \
    DP   NP
   /  \
  the NP
 /    \
 fisherman CP
      /  \
     C₂ TP
    /  \
   that DP
  /    \
  the NP CP
 /       \
  killer whale CP VP
 /         \
  Op₁ TP
 /         \
  C₁ TP
 /         \
  that V
  /       \
  VP
 /     \
  V VP
 /       \
  VP
 /   \
  V DP
 /       \
 chased the seal
```

(died)
Center-embeddings

$O_p$ c-commands $\theta'$; $O_p$ is closer to $\theta'$ than $O_p$, just as in the canonical center-embedding. The thematic operator $O_p$ cannot bind the values of the $\theta$-features, which means that the NP fisherman is not thematically bound. In addition, a licensed thematic position does not undergo Pure Merge with a nominal argument. These are local violations of the Theta Principle, which give rise to processing difficulty of the familiar type.

In the following sub-sections, I will discuss some supporting arguments to the proposed analysis in view of possible objections.

3.9. A note on binding from a Specifier

That an element in a Specifier may bind out of the phrase that it is the Specifier of in violation of the standardly assumed c-command might appear controversial. However, there is independent evidence that ‘binding out’ of the standard c-command domain is possible for Specifiers.

Reinhart (1987) points out that binding is possible even when standardly assumed c-command is not met. An example is found in (168).

(168) Every girl’s father thinks she’s a genius.

The element girl may bind she, outside of its standard c-command domain. However, under the c-command definition used in this work, girl actually does c-command she (as discussed in Kayne, 1994, p. 23-4). A question about reflexives immediately arises. In sentence (169), every girl c-commands the reflexive, and Condition A of the Canonical Binding Theory (Chomsky, 1981) is expected to be met. However, the sentence is ungrammatical.

(169) *Every girl’s father admires herself.

The ungrammaticality of (169) follows from Reflexivity Theory (Reuland, 1998, Reuland and Reinhart, 1993). The revised Condition A states that herself reflexive marks the predicate admire. Therefore the antecedent of herself can only be every girl’s father, and not every girl, since it is only the former element that is a co-argument of herself. The ungrammaticality follows then from the mismatch in features between two arguments.

Another demonstration that an element in a Specifier appears to bind outside of the standardly assumed c-command domain is the so-called crossing dependencies example (Robert Levine, personal communication):

44 ‘Standard’ c-command: a node $\alpha$ c-commands a node $\beta$ iff neither $\alpha$ nor $\beta$ dominates the other and the first branching node dominating $\alpha$ also dominates $\beta$. For an overview of the evolution of the definition of c-command, see Reinhart and Siloni (2005a).
(170) You know what Robin is like to argue with.

The element in the Specifier of TP Robin binds its copy, but does not c-command it under a standard definition. Binding is also able to cross the dependency between the element in the Specifier of CP what and its relevant copy.

Thus, even without adopting a specific c-command definition, Specifiers can bind outside. By comparison then, an operator in [Spec, CP], which is a segment of a DP, can bind out irrespective of the c-command definition. Whatever the choice of c-command may be, it should be obvious that a Specifier would always either c-command or bind out, depending on the theoretical choice that is made. Crucially, the principle for determining whether binding out is possible is whether the operator Probe occupies a Specifier.

3.10. Amelioration

Center-embeddings that contain first and second person pronouns in the subject position of the most embedded relative clause are much easier to process than sentences containing definite nouns or proper names in that position. Sentences (171) and (172) are not as difficult as (126):

(171) A book that some Italian that I have never heard of wrote will be published soon by MIT Press. (Frank, 1992)

(172) The reporter who everyone that I met trusts said the president won’t resign yet. (Bever, 1970)

Gibson (2000) reports that complexity ratings were the lowest for center-embeddings which contained first and second person pronouns in the most embedded subject position, compared to definite nouns and proper names in the same position.

Two crucial conclusions were drawn out of these data: (1) That a syntactic account of the difficulty of center-embeddings is implausible (Kluender, 1998); (2) That the type of the NP in the subject position of the most embedded clause plays a crucial part in the complexity of these sentences (Hudson, 1996).

Nonetheless, Kluender (2004) draws an explicit parallel between ameliorating factors of the complex NP island and center-embeddings of the type discussed here (what he calls center-embedding contexts). Compare:
Center-embeddings

Center-embedding

(173) ¿The woman the man the host knew brought left early.

(174) The woman someone I know brought left early.

Complex NP island

(175) ¿That’s the campaign that I finally thought of the aide who could spearhead.

(176) That’s the campaign that I finally thought of someone who could spearhead.

The substitution of a definite noun with a quantifier results in amelioration of both constructions. Why should the same factor ameliorate both types of constructions? This went unaccounted for (except for stating that these are similar processing phenomena), but is natural and in line with the observation drawn here: that center-embedded relative clauses are related to the complex NP island.45


(177) [[CP That the food [CP that John ordered] tasted good] pleased him].

(178) [[CP What the woman [CP that John married likes]] is smoked salmon].

These examples provide a clue to a possible explanation of amelioration of (171), (172), (174), (177), and (178) by asking about the syntactic status of the various embedded CPs. I now turn to explain the observed ameliorations from a syntactic point of view.

3.10.1. Amelioration by 1st and 2nd person pronouns

There is independent motivation to assume that agreement on 1st and 2nd person pronouns is subject to stricter conditions than other φ-features (such as gender and number). For instance, the 1st person pronoun in (179) may not refer to the speaker of the sentence (Kayne, 2000):

(179) The man, who is talking to you wants you to give him/*me some money.

Baker (to appear) suggests that all matrix clauses and certain embedded clauses have special null arguments generated within the CP projection designated S (for Speaker) and A (Addressee). The idea is that certain key

45 In the study of islands (Chomsky, 1977), amelioration was noted, too, but that was no reason to dismiss a syntactic account.
parameters of the ‘point of view’ aspect of the interpretation of clauses are fixed for that clause by syntactically represented elements. All uses of a first person pronoun must then be anchored by being bound by an $S$ operator, and all uses of a second person pronoun must be anchored by being bound by an $A$ operator, according to the conditions in (180).

(180) **The Person Licensing Conditions**

   a. A DP or NP $X$ is 1$^{st}$ person only if it is locally bound by the closest c-commanding $S$ or by another element that is itself 1$^{st}$ person.

   b. A DP or NP $X$ is 2$^{nd}$ person only if it is locally bound by the closest c-commanding $A$ or by another element that is itself 2$^{nd}$ person.

   c. Otherwise, $X$ is third person.

The conditions in (180) explain the fact that ordinary non-pronominal NPs are never first or second person, even when they refer to the speaker or hearer. This can be seen in (181), where ordinary DP subjects cannot trigger distinctive first or second person agreement on the copular/auxiliary verb $be$.

(181)  

   a. The man who is talking to you is/are hoping to get some money.

   b. Sorry honey, but Daddy is/are too tired to play with you tonight.

   c. Your honor was/were misinformed by the counsel for the defense. (cf. You were/were misinformed…)

It is a general fact that pronouns can be bound by operators (e.g. in resumptive pronoun constructions) but lexical NPs generally cannot be (Pronouns are bound by operators, but 1$^{st}$ and 2$^{nd}$ person pronouns are additionally bound by an $S$ or $A$ operators). This is shown by contrasts like the one in (182).

(182)  

   a. ?John$_k$, who$_k$ we all wonder whether he$_k$ will actually show up,…

   b. *John$_k$, who$_k$ we all wonder whether the boy$_k$ will actually show up,…

The conditions in (180) can also be used to explain the fact that a first person pronoun cannot refer to the speaker in (179). The sentence as a whole has an $S$ operator that refers to the speaker. What is special about it is that it also has a matrix subject that is a definite description that independently refers to the speaker, without being dependent on $S$ (and hence without being first person). Suppose that this fact is represented by giving both the operator $S$ and the matrix subject the index $<i>$. Next, suppose that the pronoun in the indirect object position of the embedded clause refers to the same individual.
If it is assigned this index, then it is bound by $S$, but it is not locally bound by $S$; the matrix subject is a closer NP that bears the same index. Therefore, the pronoun in question cannot be first person given (180a); rather, it must be third person by default:

(183) \[ \text{[CP } S \text{ [TP } \text{NP: The man who is talking to you] wants you to give him/*me, some money]} \]

In the ameliorated center-embedding cases (171) and (172), a Speaker operator may be merged with the Specifier of the relative CP.

(184) A book \[ \text{[CP } Op_k \text{ that some Italian [CP } S_m \text{ Op}_i \text{ that I have never [VP } \theta'_i \text{ heard of } \theta_i]] [\text{VP } \theta'_i \text{ wrote } \theta_i]} \] will be published soon by MIT Press.

(185) The reporter \[ \text{[CP } Op_k \text{ who everyone [CP } S_m \text{ Op}_i \text{ that I met [VP } \theta'_i \text{ trusts } \theta_i]] [\text{VP } \theta'_i \text{ wrote } \theta_i]} \] said the president won’t resign yet.

Since $S_m$ is a head of a functional projection, the latter dominates the projection of the thematic operator. Therefore, $Op_i$ cannot bind out and is no longer closer to the Goal $\theta'_i$ than $Op_k$. Intervention is circumvented. No violations of the Theta Principle emanate and the sentences are correspondingly easy to process in accordance with actual performance. The same account can be applied to 2nd person pronouns by replacing $S$ with $A$.46

Note that the insertion of an operator as a head functional projection is not an ad-hoc explanation of amelioration. The evidence for $S$ and $A$ operators are independently motivated.

Beyond the suggestion of $S$ and $A$ operators, there is good evidence that quantifiers, which head relative clauses as in (171) and (172) may contribute to amelioration, too. When a head is quantificational, the quantifier is not included in the noun phrase inside the relative clause (Stockwell, Schachter and Hall-Partee, 1973). For instance, (186) does not entail (187):

(186) All the boys who left early missed the fun.

(187) All the boys left early.

In examples (174) and (176) that involve quantifier substitution (and similarly also in (171) and (172), which have quantifiers as head nouns), the projection of the quantifier dominates the thematic operator (as in the cases

46 $S$ and $A$ are quantificational operators that do not necessarily occupy the edge of CP. When a certain XP is in an object position, they would reside in the Spec of VP, when in subject position, in the Spec of CP. This is why the subject position, [Spec, TP], of center-embedded constructions causes amelioration, when occupied by 1st and 2nd person pronouns. Note that the subject position has temporal properties, and the subject is an anchor for an event. That is why it is different from object positions.
of the Speaker and Addressee operators). The projection of the quantificational head thus circumvents any illicit Probe-Goal relations. As a result, it is not possible to violate the Theta Principle, and no processing difficulty is experienced.

3.10.2. Amelioration of doubly embedded free-relatives and sentential subjects

The ameliorated examples (177) and (178) can be similarly explained by inquiring about the Probe-Goal relations created during the parsing of these structures and their relevant syntactic properties.

It is a syntactic fact that that the food that John ordered tasted good in (177) is a sentential subject, and that what the woman that John married likes in (178) is a free relative clause. The clausal heads of free relatives or sentential subjects, in contrast to (restrictive) relative clauses, do not possess the ability to Attract (values of features of) thematic roles, as they do not contain a thematically licensed position that requires Pure Merge with an argument (no 0-operator).

As can be seen from Tree 18 of sentence (177), there is no hope for an intervention effect. The operator $Op_i$ cannot bind out, since it is dominated by the $C_1$-head projection. Consequently, intervention is circumvented. No violation of the Theta Principle occurs, and the sentence is easily comprehended.
Similarly, consider Tree 19 of sentence (178).
There is no thematic Goal in the Spec of VP *like*. If Op were to bind out, there would be no values of 0-features in the Specifier position of the VP *likes* and therefore no matching Goal in a Specifier position, which it could bind. The $C_1$-head does not Attract the values of the features of a thematic role, simply because it is a head of a free relative clause and not of a relative NP. There is no intervention and the sentence is predicted to be easily processed in accordance with actual performance.

The thematic interpretation of the free relative occurs during parsing only when the parser arrives at *is*, when it is possible to license the relevant thematic roles of *like*. Thus, there is a distinction between a syntactic requirement that values of 0-features be merged with an edge position, as in restrictive relative clauses, and a non-syntactic, semantic requirement on binding of arguments, as in the case of free relatives.

Examples (177) and (178) provide a strong motivation for considering the delicate syntactic Probe-Goal relations formed during parsing. They are also striking examples of the necessity to provide a syntactic account of amelioration and for processing difficulty of center-embeddings.

The conclusion from this discussion is straightforward. Amelioration begs for a syntactic account of various center-embedded constructions contra to what has been commonly thought of these constructions. It emphasizes the viability of a syntactic account, along with its tight relation to syntactically restricted relations formed during parsing.

### 3.11. Center-embeddings in left-branching languages

A well-known claim among linguists is that speakers of left-branching languages can process three levels of center-embedding, finding two levels of center-embedding completely comprehensible, whereas speakers of right-branching languages consume their full processing capacity already at two levels of center-embeddings. This claim can be traced back to Rosenbaum and Kim (1976) who compared the processing difficulty of center-embeddings between left- and right-branching languages. Specifically, they compared the comprehension of English and Korean center-embedded object relatives. Rosenbaum and Kim argued that although the sentences were relatively difficult for native speakers of Korean, they found the sentence to be comprehensible. If the structures of center-embeddings in both types of language are symmetric, then in a syntactic approach to language, both should yield an intervention effect and produce processing difficulty, all other factors beings equal. How can this, if true, be accounted for by the current approach?

Consider the parsing steps of center-embedded doubly-relativized clauses in Korean, Rosenbaum and Kim’s critical example:
The stages of the parse are as follows. The three Nominative-marked DPs are entered to the parse. When _love_ is encountered, it is identified as an obligatorily transitive verb, and its maximal theta grid is retrieved to have an external and internal theta-role. The _student_ is pMerged with the external thematic position being marked Nominative. Although all DP’s are marked as potential subjects of _loved_, only the most local attachment is executed (a limitation constrained perhaps by Adjacency or Pure Merge). The next constituent is _girl-Dat_. Abiding to Adjacency, an argument NP that may be pMerged in the internal thematic position of _love_ will not appear (the internal thematic position is to the left of _love_), such that the internal theta-role licenses with the direct object position. A relative clause must be built, as also indicated by a post-verbal NP, in accordance with grammar.47 The C-head attracts the values of θ-features to the Specifier of the VP allowing for the binding by a thematic operator under Probe-Goal relations.

The relative NP in turn needs to be pMerged with a thematic position and is merged in turn with the thematic position of the preposition _for_ (in actuality, the Dative marker). The next constituent is _buy_, identified as an obligatorily transitive verb and retrieved to have two theta-roles in its maximal theta grid. The Nominative-marked _our father_ is pMerged with the external theta-role of _buy_. An argument for the internal theta-role of _buy_ is not going to appear in accord with Adjacency, and so, the internal theta-role is pMerged with the direct object position. The next constituent is _the vase_, and since it is a post-verbal NP, a relative clause is formed. As before, the local C-head attracts the thematic Goal to a Specifier to allow binding by a thematic operator. This is the crucial situation of the parse given in Tree 20 (omitting material that is still to appear):

---

47 In Korean/Japanese the head noun follows the relative clause.
There is no possible intervention effect. \( Op_j \) c-commands \( Op_i \), and is closer to its feature matching Goal \( \theta_j \) than \( Op_i \) as required. No local violation of the Theta Principle is generated and the sentence is consequently rendered relatively unproblematic to comprehend.
The Korean sentence differs from the canonical English (126), since two elements (the direct object and indirect object of *buy*) are licensed by the same verb (*buy*), whereas in English the relativized NPs are licensed by two different verbs.

The conclusion is that Rosenbaum and Kim compared two structurally different sentences. Superficially, from looking at the linear order of the words in Korean, sentence (188) has center-embedded relative clauses and resembles the canonical English example. However, the Korean sentence is sequentially (incrementally) parsed, closing one relative clause after the other, much like unproblematic sequential English equivalents (for instance, *the dog chased the cat that bit the rat that died*).

In fact, it is impossible to construct an equivalent structure to the English (162) in Korean, or in any strictly left-branching language (such as Japanese, Turkish, Assamese), due to the structure of relative clauses where the head-noun must follow the relative clause. Any comparison of the sort that Rosenbaum and Kim attempted to make actually referred to the linear sequence of constituents (though in compliance with the definition of self-embeddings), not to hierarchical structure or to parsing considerations. It follows that their claim that speakers of Korean can process one more level of embedding than speakers of English was misconceived; they compared two different structures, the parsing of which gives rise to different temporal representations. There is no reason to assume then that there is a mysterious difference in short-term memory between speakers of the two kinds of languages. Speakers of both types of language can easily process consecutive order of constituents. English speakers encounter a difficulty in sentences such as (162) caused by the local structural-hierarchical relations.

The examples in English and Korean provide strong evidence to the claim that it is essential to follow the temporal steps of sentences and consider the hierarchical syntactic relations formed incrementally. Further, they augment the argument that in both left and right-branching languages, Theta Attachment is a crucial heuristics for determination of potential attachments.

### 3.12. Center-embeddings of complement clauses

Center-embeddings of complement clauses in left- and right-branching languages inflict milder processing difficulty than the English canonical examples. The following are examples of center-embedded complement

---

48 An informal survey among linguists, whether center-embedded complex NP islands can be constructed in Japanese, Turkish and Assamese, all left-branching languages, revealed that they were indeed impossible to construct (Thanks to Reiko Vermeulen, Cem Keskin and Shakuntala Mahanta).
clauses in two left-branching languages, which are easy (from Hagstrom and Rhee, 1997):

(189) **Japanese**

Haha-ga titi-ga hukigenna akatyan-ga
Mother-Nom father-Nom fussy baby-Nom
naita to itta to omotteiru
cried that said that thinks

‘My mother thinks that my father said that the fussy baby cried.’

(190) **Korean**

Sunhi-un Chelswu-ka Yenghi-ka uless-tako
sunhi-Top Chelswu-Nom Yenghi-Nom cried-that
malhayss-tako mitmunta
said-that believes

‘Sunhi believes that Chelswu said that Yenghi cried.’

The Japanese and Korean sentences can be more readily comprehended than canonical center-embeddings in English. Uehara (2003), too, found that Japanese speakers were successful at comprehending Japanese center-embeddings of this type, despite the fact that they were more difficult to process. It appears that Japanese is more tolerant of embedding than English. The difference in the quality of the difficulty and the higher comprehension ability in Japanese should be explained relying on structure.

Complement clauses are standardly assumed to be arguments selected by the head noun (and not adjoined, like relative clauses). However, I will claim that complement clauses selected by certain head nouns (‘factive’ nouns such as *fact, report,* etc.) are like appositions (non-restrictive) and for that reason are adjoined to the head noun, relying on Kayne (1994). The relation between the head noun and the complement clause is expressed by a semantic operator. The operator is merged with the Specifier position of the CP, much like in restrictive relative clauses. Consider, for instance, the English sentence (191).

(191) ¿The finding that the disclosure that the report was true was outrageous was alarming.

As Tree 21 below of sentence (191) demonstrates, there is no possibility of an intervention effect; the operator *Op,* does not have a thematic Goal in an edge position.
Since no intervention arises, the Theta Principle is not locally violated, and
the sentence does not incur difficulty. However, $Op_i$ is still found in the
search domain of $Op_k$, and may therefore cause disturbance (but, obviously,
there is no intervention) in the formation of the relation between $Op_k$ and its
relevant clause. The disturbance is semantic in nature: the operators are
assumed on semantic grounds. Further, there are no Attract operations.
Semantic disturbance is found in other English constructions, too. The following example is of a relative NP embedded in a tough construction:49

(192) Dan is too stubborn [Opₜ [PRON to expect people [Op_i that hardly \theta_i understand the situation] to talk to \theta_i]].

The example is relatively easily processed but is predicted to be difficult, since there could be an intervention effect between the Probe Opₜ and its Goal \theta_i by Opᵢ.

However, Opₜ is dominated by a non-relativizing CP. Under a Probe-Goal analysis, the C-head does not Attract the thematic values associated with \theta_i to the edge – only a C-head of a relative NP does that. Thus, \theta_i remains in-situ and its thematic properties are semantically associated with Opᵢ; indeed Opₜ is a non-thematic operator assumed to be necessary on semantic grounds. The Merge of Opᵢ causes disturbance, but it does not lead to an intervention effect. The thematic Probe Opᵢ cannot bind \theta_i, because the requirement is that a thematic goal be in a Specifier position. No binding that may lead to intervention occurs, and the sentence is easier to process than canonical center-embeddings.

In the Japanese and Korean examples (189), (190), respectively, there is no hope for an intervention effect or a disturbance as Tree 22 for Korean demonstrates:

---

49 I thank Tali Siloni from Tel-Aviv University for pointing out this example to me.
3.13. Center-embeddings of relative clauses and complement clauses

When a relative clause is embedded in a complement clause, the resulting sentence is relatively easily processed and comprehensible (sentence (193) below). In contrast, when a complement clause is embedded in a relative clause, as in (194), the sentence turns out to be incomprehensible (examples taken from Gibson (2000)):

Since verbs select sentential complements (in English, the latter are selected by a special group of nouns), the intervention/disturbance situations that arise in (191) and the canonical (162) simply do not arise. Hence, (189) and (190) are processed more easily, in accord with actual performance.  

Embedding (190) in another layer results in total incomprehensibility (Hagstrom and Rhee, 1997). This might reflect an upper limit on the ability to compute thematic relations during parsing.
Relative clause embedded in a complement clause

(193) The fact that the employee who the manager hired stole office supplies worried the executive.

Complement clause embedded in a relative clause

(194) The executive who the fact that the employee stole office supplies worried hired the manager.

How can this contrast be explained under the thematic parser purported here? Following the relevant parsing stages of (193) from the point stole is encountered reveals that by Adjacency, a theta-role is pMerged to license the direct object position of hired (an obligatorily transitive verb). The relevant Attract operation by the relativizing C1-head is performed by merging the theta-role to the Spec of VP to enable thematic binding. The theta grid of stole was retrieved, too, and it pMerged the relative NP the employee the manager hired with its subject position. A theta-role is discharged to license the argument office supplies when it is encountered, as Tree 23 below illustrates.

Tree 23
The $C_2$-head is non-relativizing, and no (values of $\theta$-features) of theta-roles are attracted to the Specifier position of *stole*. Therefore, there is no possibility for an intervention effect (although there is disturbance): there is no Goal for $Op_i$ (the complementizer *who*), similarly to embedded sentential subjects and free relatives.

In the case of (194), at the relevant stage when *hired* is encountered (presented in Tree 24 below), the internal theta-role of *worried* can be licensed. The (values of the $\theta$-features of the) theta-role is attracted by $C_2$, a thematic relativizing head, to a Specifier position.

Tree 24
An intervention between $Op_3$ and $\theta'$ emanates by $Op_2$, preventing the former from being bound. The Theta Principle is locally dissatisfied in the familiar fashion, and the sentence is rendered difficult to process.

Example (194) further shows that only thematic relativizing heads may Attract (values of the $\theta$-features of) theta-roles, and that appropriate distinctions among various heads of clauses must be made.

3.14. Center-embeddings and adjuncts

The question this section deals with is the effect of constituents that bear a theta-role, yet are not lexically selected by a predicate (adjuncts), on the degree of difficulty experienced in center-embedded constructions. Consider (195) taken from Reinhart (1987):

(195) $\exists$Max $[VP_1$ promised to $[VP_2$ persuade Bill to $[VP_3$ go to Chicago on Monday] on Tuesday] yesterday].

When the temporal factor of this sentence is not considered as in the study of competence, the sentence conforms to the definition of center-embeddings (self-embedding, too), as it contains verb phrases that are center-embedded in one another.

These structures are supposed to incur difficulty, like canonical center-embeddings. However, under a parsing account, where the temporal and lexical aspects are taken into account, no intervention is expected. The sequence of VP's (and NPs) requires the nesting of each incoming VP in the former; for instance, persuade selects a CP that dominates VP$_3$, as much as promise selects VP$_2$ (a CP). The added complexity of this structure is felt only when the temporal adjuncts appear and is due to the inability of the parser, being motivated by Theta Attachment, to decide to which VP to attach them. This situation is similar to theta-role indeterminability situations (Section 2.1), where Theta Attachment is unable to adjudicate between attachments. The same sentence without adjuncts does not cause difficulty:

(196) $\forall$Max promised to persuade Bill to go to Chicago.

All verbs may potentially license all adjuncts in (195); there is no lexical information on the basis of which the parser may decide on the correct attachments. The same approach is valid for the following center-embedded NPs:

(197) $\exists$Of particular interest was $[NP_1$ an excerpt from $[NP_2$ an old book about $[NP_3$ Jewish prayers] in the Oxford library] in the new reader].

Although the sentence does not have verbs making thematic relations harder to investigate, the same type of puzzlement for the parser occurs here, too. It is not clear to which NP each locative adjunct must be attached.
These examples provide a good reason for the inclusion of the factor of time in the study of syntax, and emphasize the important difference between grammatical competence and parsing performance approaches to language study.

3.15. A detour to subject- and object relative clauses

A natural upshot of the approach defended here is the explanation of the difference in processing between subject- and object relative clauses. It is an established fact that object relative clauses are ‘more complex’ than subject relative clauses, in English and other languages (for Chinese: Lin, 2006, Lin and Bever, 2006, for Japanese: Miyamoto and Nakamura, 2003). For example, (198) is more complex than (199):

(198) The tourist [that the lion attacked] was already wounded.
(199) The lion [that attacked the tourist] was already wounded.

The greater complexity of object relative clauses compared to subject relative clauses was detected by various experimental modalities, for instance, tracking of eye-movements and self-paced reading times (Hakes, Evans and Brannon, 1976, Holmes and O’Regan, 1981, King and Just, 1991, Ueno and Garnsey, 2005, Waters, Caplan and Hildebrandt, 1987). Note that this effect cannot be driven by lexical frequencies or real world plausibility since these are controlled for between the two sentences; or discourse context. The reason for the effect must reside in structure. This fact did not receive a proper grammatical explanation (it was regarded as a performance issue) up until the advent of the Minimalist Program and its Strong Empirical Hypothesis.

Within the thematic parser, the difference between (198) and (199) is straightforward. In (198), the direct object thematic position is pMerged with a theta-role. The values of 0-features of the theta-role must be attracted to an edge, the Specifier position of the relevant VP. In (199), the verb thematically licenses the Specifier position of the TP and no Attract operation is required. In the object relative clause (198), the Attract operation is obligatory, consuming more cognitive resources manifested by higher relative difficulty.

3.16. Conclusion: the difficulty of center-embeddings as local ungrammaticality

A new realization has presented itself by considering canonical center-embeddings with the Thematic Transparent Model. Viewing these constructions as closely related to islands has lead to the suspicion that the
difficulty manifested in them is due to ungrammaticality, as it is the case in any other islands.

It was necessary first to exclude other factors that have been suggested in the literature that incur difficulty in center-embeddings. I found that the factors of number of self-embeddings, or the number of self-embeddings and object relativizations did not explain gradation of difficulty of various self-embeddings, since those were incompatible with data obtained in a Magnitude Estimation experiment. Furthermore, other suggestions as to the difficulty of center-embeddings, e.g. the non-local nature of discourse structure in center-embeddings or the effect of assuming potential similar syntactic roles, did not comply with the results of a Magnitude Estimation experiment conducted to test these hypotheses. Additionally, severe doubts have been provided that question the validity of such claims about the origin of the difficulty of center-embeddings. The results of both experiments were compatible with the observation that the difficulty of center-embeddings is related to their status as instances of islands.

Following the parsing steps as guided by Theta Attachment has led to the surprising finding that an intervention effect in canonical center-embeddings emanates during parsing, making the local satisfaction of the Theta Principle impossible. In turn, it revealed why other nested constructions (free relatives, sentential subjects, etc.) do not induce difficulty and how amelioration by 1st and 2nd person pronouns can be explained on syntactic grounds. The model satisfies (119), thus providing a complete model of parsing performance. The ability to successfully generalize the model to data it was not meant to explain provides strong support to its 'goodness of fit'.

Vosse and Kempen (2000) suggest a psycholinguistic model that complies with conditions (119)-(123), too. The model, at its core, is based on the lateral inhibition between incompatible attachment alternatives. It does not construct syntactic trees. Instead, it creates a network of connections between nodes of basic syntactic heads (N, D, P, V, etc.).

Despite its satisfaction of parsing adequacy, the complexity of the model itself is so high, as a quick glance of an example of inhibitory links illustrates (Figure 5), that a computer is required to implement it.
And as Pitt and Navarro (2005) point out, when a model is so complex, it may run the risk of having very low explanatory power in the sense that a generalization is not possible. Applying Ockham’s razor, the Thematic Transparent Model should be preferred.

In essence, the intervention effect found in canonical center-embeddings is a manifestation of the parser’s inability to overcome the local violation of the Theta Principle: an identified argument cannot be associated with a thematic position, and a thematic position does not Merge with an argument. This situation explains the relative lighter processing difficulty of garden paths relative to center-embeddings. In garden paths, the violation of the Theta Principle concerns only the failure of the parser to pMerge an argument with a licensed thematic position or vice versa. The parser is punished for its adherence to Theta Attachment (when the repair operation is raising, violating the OLLC). In center-embeddings, it is impossible to reject Theta Attachment, because of intervention. This directly accounts for the elevated difficulty of center-embeddings compared to garden paths. We may now see in Flowchart 3 below how the Thematic Transparent Model accounts for the various kinds of difficulty.
Flowchart 3: Deciding when a garden path effect and a center-embedding effect are obtained

The thematic parser unites two seemingly disparate processing phenomena, and provides them with a single cause.

Insofar as processing difficulties manifested in center-embeddings and garden paths is a psychologically real phenomenon (difficulty is a measurable effect of the brain), it can be said, with even greater certainty
now, that violations of the Theta Principle correspond with these difficulties. The Theta Principle is represented in the mind/brain and is therefore psychologically real.

Beyond the purposes of this inquiry, the discovery that center-embeddings are related to islands and that the difficulty associated with them is intricately related to the temporal local application of grammar, sheds new light on the human mind and on the nature of these constructions. First, it makes center-embeddings of no special status: they conform to the same parsing/grammatical principles as any other linguistic phenomenon. This stands in stark contrast to what has been thought about these constructions for the past fifty years: that they are immune to a syntactic account. Processing difficulty sensed in center-embeddings is local grammaticality judgment and is therefore a direct reflection of grammar.

Second, canonical center-embeddings can no longer serve as an example of the limited capacity of short-term memory. Center-embeddings are within the capacity of short-term memory. The difficulty is not due to a memory limitation on the parser, but to the inability of comprehension as manifested by the local dissatisfaction of the Theta Principle.

Short-term memory is, uncontroversially, a finite automaton that allows for the operations of the parser. However, from the discussion here, it appears that the operations that the parser executes are not dependent on the limitations of a finite automaton. In essence, short-term memory is domain-general, and not parser-specific. What center-embeddings reflect is a local violation of the Theta Principle, which is a fundamental and unique grammatical principle of human language. Its dissatisfaction is simply a paraphrase of the failure of producing a full proposition. This means that the structure the parser builds incrementally cannot be interpreted by the Conceptual-Intentional (C-I) interface; that there is a failure of comprehension.
I posed a question in the first chapter, which served as motivation for this investigation:

(200) Can the study of parsing performance provide evidence that a principle of the theory of Universal Grammar is psychologically real?

I contended that an argument for the psychological reality of a certain global grammatical principle can be put forward if a model which incorporates this principle has ‘goodness of fit’. To accomplish this, I hypothesized that the human sentence processing mechanism is transparent: the core of syntactic parsing consists of the local application of global grammatical principles.

Chapter 1 revalidated the Thematic Transparent Model as a model of the human sentence processing mechanism by accounting for alleged counterevidence mostly found in psycholinguistic literature. I argued that the Thematic Transparent Model parallels the processing difficulty experienced in garden path constructions with a local violation of a global grammatical principle: the Theta Principle.

Chapter 2 had two aims. The first was to establish the ‘goodness of fit’ of the Thematic Transparent Model with experimental evidence. The model predicts random occurrence of the garden path effect in theta-role surplus sentences. Alternatively, performance on such sentences should differ from performance on standard garden path sentences. These sentences are predicted to demonstrate such random performance because the guiding principle of the parser, Theta Attachment, does not adjudicate. I presented
the predictions on theta-role surplus sentences of various other models. All
other models make predictions that differ from the thematic parser.

I found that in Dutch versions of theta-role surplus sentences, the
occurrence of the garden path effect was at random and that performance on
such sentences differed from performance on standard garden paths. The
chapter therefore supplied empirical evidence that the thematic parser
corresponds with human parsing performance on theta-role surplus
sentences, and that other models are incompatible with these data. The
chapter established Theta Attachment as a guiding principle of syntactic
attachments, and the argument that human sentence processing is driven by
thematic considerations.

The second aim of this chapter was to establish the validity of the
experimental paradigm that was adopted for the purposes at hand: Magnitude Estimation. Magnitude Estimation measures subtle gradation in
grammaticality judgments. I chose the method on the basis of the argument
that the garden path effect is a local judgment of ungrammaticality. I
compared the results of Magnitude Estimation experiment with those of an
eye-movement experiment. I was able to establish that Magnitude
Estimation is a sensitive method, which not only replicated the results of the
eye-movement experiment, but also outperformed it. The method is
adequately suitable for the study of parsing performance.

To explore the possibility that the Thematic Transparent Model has
more empirical coverage than just the prediction of parsing performance on
garden path constructions, I generalized the model to center-embeddings in
Chapter 3.

I observed that the processing difficulty of center-embeddings could
be related to their status as islands. I conducted experiments specifically
designed to test various hypotheses of other models concerning center-
embeddings, and excluded these models as possible explanations. The failing
of other models to account for the difficulty in center-embeddings is partially
due to their failure to observe the fact that center-embeddings incorporate
islands.

I proceeded to argue that the processing difficulty of center
embeddings is a manifestation of the local violation of the Theta Principle,
and accounted for various types of ameliorations reported in the literature. I
also considered center-embeddings in left-branching languages and
accounted for them within the thematic parser.

The success of explaining the processing difficulty of center-
embeddings by grammatical means within the Thematic Transparent Model
achieves three goals. First, the model fully satisfies requirements of parsing
adequacy and can be considered a complete model of human sentence
parsing, for both left- and right-branching languages. Second, it establishes
the model’s reasonable ‘goodness of fit’, in the sense that the model explains
the data it was supposed to explain and another body of data. Third, it
provides firmer grounds to the argument that the Theta Principle is
psychologically real, and to the fact that human sentence parsing is guided
by Theta Attachment.

It is the prediction of the Thematic Transparent Model that canonical
center-embeddings will invoke the most costly processing difficulty
phenomena: they contain local violations of the Theta Principle, augmented
by two Attract operations. The satisfaction of the Theta Principle is
prevented by an intervening element.

Garden path constructions, which have a local violation of the Theta
Principle, but no interference, invoke a lesser degree of difficulty.

Processing phenomena (garden paths, and center-embeddings) which
seem disparate are in fact multifaceted manifestations of a single principle,
the Theta Principle. The classification of the various processing phenomena
as a violation of the Theta Principle is illustrated in Flowchart 4.
Flowchart 4: Deciding which processing effect the violation of the Theta Principle leads to, and the type of violation

Possible violations of the Theta Principle are diagnosed: A theta-role has no argument to pMerge with, and/or an argument has no theta-role with which it may pMerge. The parser must decide whether there is a local grammatical possibility to cure the local violation of the Theta Principle. If Adjacency is relevant, then the local violation is repaired without cost.
If the cure involves the rejection of Theta Attachment (raising, which violates the OLLC), then the local violation of the Theta Principle causes processing difficulty of the garden path type.

Center-embeddings involve a situation where a theta-role is without an argument and an argument is without a theta-role. Interference does not permit rejection of Theta Attachment or the satisfaction of the Theta Principle (which is not the case in garden paths and therefore the difficulty is less severe). Therefore, the violations lead to the effect of the type found in center-embeddings. Otherwise, the violations are not costly.

Following Pritchett (1992a), I have shown that an approach to parsing formulated directly in terms of grammatical principles is cognitively sufficient. The success of the model in covering a wide body of compelling data from right- and left-branching languages strongly suggests that the human natural language processing mechanism is universal, and may be characterizable wholly in terms of grammar, where the grammar is viewed not as a system of rules, but as a set of conditions on representation, and where the parser applies the principles of grammar locally at every point during the processing of the input string:

“More broadly, both processability and grammaticality are accounted for by the same principles, though in somewhat different fashions. Ungrammaticality results from the global violation of some grammatical constraint. Unprocessability, on the other hand, is attributable to certain local violations of grammatical principles. Success in accounting for an extremely wide range of processing phenomena in a simple and unified fashion both in English and across typologically distinct languages has provided strong evidence that the core of Parsing theory is derived from the theory of Grammar.” (Pritchett, 1992a, pp. 155-156)

The empirical evidence in favor of the model’s ‘goodness of fit’ and its success in generalizations converge to provide a single answer to (200): the Theta Principle is psychologically real.

Failure to satisfy the Theta Principle corresponds to external reality – a physically real phenomenon in the human brain – manifested by processing difficulty. The effect is measurable, and is ‘externally real’. By correspondence, the Theta Principle is a real property of the human brain.

The Theta Principle is not only a grammatical principle vital for parsing. It plays a crucial role in language acquisition. The principle is directly linked (via Theta Attachment) to argument structure. The acquisition of argument structure plays a crucial role in the cognitive development of children. Within the Minimalist Program, the acquisition of argument
structure is part of building up of the lexicon (Chomsky, 1995, Section 1.2, p. 30). Even in so-called usage-based theories, this much is agreed upon; proponents of construction grammar emphasize the importance of the acquisition of argument structure in language acquisition (Goldberg, 1995, Tomasello, 2003).

Any human child must grow the fundamental cognitive ability to discern arguments out of the input, distinguish between them and develop the structural implications that follow (the fundamental distinction, e.g., between arguments and adjuncts). Human beings are predisposed to determine argument structure in the input, a tendency expressed by Theta Attachment. Much research is still required to crystallize the connection between the stages of acquisition of argument structure and grammatical maturation, coupled with the development of parsing capabilities. This appears to be a promising avenue for future research.

There is compelling evidence that suggest that the Theta Principle is innate insofar as the Theta Principle implies knowledge of labels of thematic roles. Goldin-Meadow (2005) studied the spontaneous gestures of deaf children (i.e. deaf children obviously not exposed to spoken language but also not to sign language). Their gestures were taken to represent a pre-linguistic stage. She was able to show that such children consistently gestured for arguments as a function of the labels of thematic roles. This suggests that the labels of thematic roles are innate (resilient, in her terms).

Indeed it must be the case that the Theta Principle is part of human beings’ biological endowment. If the Faculty of Language in the Narrow sense (FLN) comprises of:

“… only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces” (Hauser, Chomsky and Fitch, 2002, p. 1573)

then FLN must contain the Theta Principle, since it is an interface condition that the organ must satisfy for it to function at all (as in Chomsky, 2004, p. 106). It is a necessary interface condition between the computational system of human language (CHL) and the conceptual-intentional (C-I) interface.

The Theta Principle as such expresses a fundamental property of the Human Mind. It expresses human beings’ conceptualizations and their relation to grammar. By saying that the parser must satisfy this principle, it is meant that a well-formed proposition makes sense; that it conforms to human beings’ linguistic coding of the way they perceive the world around them. As such, it is one of the most fundamental properties of Human Language and indeed at the core of the Human Mind.
Appendixes

Appendix I

List of test sentences for experiment 1

The first sentence in each quartet is Minimal Attachment-long version, the second Non-Minimal Attachment-long version, the third Minimal Attachment-short version, the fourth the Non-Minimal Attachment-short version.

1.  a. Sally was relieved when she found out the answer to the difficult physics problem.
 b. Sally found out the answer to the difficult physics problem was in the book.
 c. Sally was relieved when she found out the answer.
 d. Sally found out the answer was in the book.

2.  a. It’s possible the detective didn’t see the man with a gun in his hand.
 b. The detective probably saw the man with a gun in his hand fall over.
 c. It’s possible the detective didn’t see the man.
 d. The detective probably saw the man fall over.

3.  a. I wonder if Tom heard the latest gossip about the new neighbors.
 b. Tom heard the latest gossip about the new neighbors wasn’t true.
 c. I wonder if Tom heard the gossip.
 d. Tom heard the gossip wasn’t true.

4.  a. It appears that Sherlock Holmes didn’t suspect the very beautiful young countess.
 b. Sherlock Holmes didn’t suspect the very beautiful young countess was a fraud.
 c. It appears that Sherlock Holmes didn’t suspect the countess.
 d. Sherlock Holmes didn’t suspect the countess was a fraud.

5.  a. I suppose it’s possible that the clerk knows the woman wearing that outrageous peacock hat.
 b. I suppose that the clerk knows the woman wearing that outrageous peacock hat is crazy.
 c. I supposed it’s possible that the clerk knows those women.
 d. I suppose that the clerk knows those women are crazy.

6.  a. Nobody knew why the speaker concluded his very interesting but technical lecture with such haste.
 b. Apparently the speaker concluded his very interesting but technical lecture has not been a success.
 c. Nobody knew why the speaker concluded his lecture with such haste.
 d. Apparently the speaker concluded his lecture had not been a success.
7.  a. The commissioner suspects that the financial committee failed to mention the very large accounting error.
   b. The financial committee failed to mention the very large accounting error was their own fault.
   c. The commissioner suspects that the financial committee failed to mention the error.
   d. The financial committee failed to mention the error was their own fault.
8.  a. They say the city council argued the radical young mayor’s position forcefully.
   b. The new city council argued the radical young mayor’s position was immoral.
   c. They say the city council argued their position forcefully.
   d. The new city council argued their position was immoral.
9.  a. The lawyers think his second wife will claim the entire family inheritance.
   b. His second wife will claim the entire family inheritance belongs to her.
   c. The lawyers think his second wife will claim the inheritance.
   d. His second wife will claim the inheritance belongs to her.
10. a. John agreed that politicians typically don’t explicitly announce their plans and programs.
   b. Politicians typically don’t explicitly announce their plans and programs are completely unrealistic.
   c. John agreed that politicians typically don’t explicitly announce their programs.
   d. Politicians typically don’t explicitly announce their programs are completely unrealistic.
11. a. After she’d had dozens of tests I think Julia finally believed the unconventional young doctor’s diagnosis.
    b. After she’d had dozens of tests Julia finally believed the unconventional young doctor’s diagnosis was accurate.
    c. After she’d had dozens of tests I think Julia finally believed the diagnosis.
    d. After she’d had dozens of tests Julia finally believed the diagnosis was accurate.
12. a. We were surprised that Sam remembered our fiftieth wedding anniversary.
    b. Sam remembered our fiftieth wedding anniversary would be quite soon.
    c. We were surprised that Sam remembered our anniversary.
    d. Sam remembered our anniversary would be quite soon.
13. a. We figured that Tom probably forgot most of his camping equipment.
    b. Tom probably forgot most of his camping equipment had been stolen.
    c. We figured that Tom probably forgot the flashlight.
    d. Tom probably forgot the flashlight had been stolen.
14. a. Nobody realized that the policeman immediately recognized all of the people in the car.
   b. The policeman immediately recognized all of the people in the car were completely drunk.
   c. Nobody realized that the policeman immediately recognized the thieves.
   d. The policeman immediately recognized the thieves were completely drunk.

15. a. Everyone thinks the democrats first suggested a new health insurance program.
   b. The democrats first suggested a new health insurance program was needed.
   c. Everyone thinks the democrats first suggested this program.
   d. The democrats first suggested this program was needed.

16. a. James was pleased that the press reported the entire sordid affair.
   b. The press reported the entire sordid affair began as a prank.
   c. James was pleased that the press reported the incident.
   d. The press reported the incident began as a prank.
Appendix II

Instructions to subjects

You are about to participate in a psycholinguistic experiment that tests the complexity of sentences, that is how difficult you find sentences to understand, or how long it takes you to comprehend a sentence.

The experiment is intended ONLY for native speakers of English, and takes about 20 minutes.

The experiment consists of three parts. In the first part, you will rate the lengths of lines, as practice. First, a reference line will appear. Give it any positive number which reflects its length, in your opinion. For instance, if you think the reference line is 20mm in length, rate it with the number 20.

Next, more lines will appear. You will have to rate the lengths of these lines in proportion to the reference line. For example, if a line seems 3 times longer than the reference line, rate it 60. If it appears 4 times shorter, rate the line with a 5.

In the second part, you will practice the rating of the complexity of a few sentences in comparison to a reference sentence. The principle is the same as in the line part; only instead of lines, there are sentences now.

First, you will assign a reference sentence with any positive number you like, corresponding to the sentence complexity (in your opinion). Let’s say you choose to rate the reference sentence with 300. After you’ve done that, you will rate the complexity of the subsequent stimulus sentences in proportion to the reference sentence.

So if a stimulus sentence is two times less complex than the reference sentence, rate it 150. If a stimulus sentence seems 4 times more complex than the reference sentence, then rate it 1200. Trust your intuitions about the complexities of the sentences.

After each sentence, you will have to answer a short comprehension question about the sentence, and the computer will let you know whether your answer is correct or not.

It will not be possible to change your reply once you’ve given it, so make sure you understand the sentence, and can answer correctly!

The third and last part is the experimental part and there you will have to rate sentences in proportion to a reference sentence and answer questions about the sentences, just like you did in the second part.

Please note that some of the sentence may appear without proper punctuation, to be nonsensical, or simply wrong. This is part of the experiment, so try to do your best.

PLEASE KEEP IN MIND:

* To judge each sentence IN PROPORTION to the reference sentence.
* NOT to use a zero as the reference rating.
* To use ANY POSITIVE NUMBER you think is appropriate.
* To use HIGH numbers for ‘complex’ sentences, LOW numbers for ‘simple’ sentences and INTERMEDIATE numbers for sentences, which are between ‘complex’ and ‘easy’.
* To try to use a WIDE RANGE of numbers and to distinguish as many degrees of complexity as possible.
* To try to make up your mind QUICKLY, NEVER spend more than 10 seconds on a sentence, and base your judgments on your first impressions.
Appendix III
Detailed instructions

Thanks for choosing to participate in this experiment. The experiment consists of several parts. Initially, you will have to enter personal details. This information will be treated confidentially, and will not be made available to a third party. None of the responses collected in this experiment will be associated with your name in any way.

If you have any questions about this practice, please contact the experimenter.

Please be careful to fill in the personal details correctly, as otherwise we will have to discard your responses. We ask you to supply the following information: your name and email address; your age and gender; whether you are a linguist or not, and your mother tongue.

After pressing the BEGIN button, the experiment will start. The experiment will consist of the following three parts:

1. Training session: Judging line length
2. Practice session: Judging sentences
3. Experiment session: Judging sentences

Part 1: Judging Line Length

Before doing the main part of the experiment, you will do a short task involving judging line length. A series of lines of different length will be presented on the screen. Your task is to estimate how long they seem by assigning numbers to them. You are supposed to make your estimates relative to the first line you will see, your reference line. Give it any number that seems appropriate to you, bearing in mind that some of the lines will be longer than the reference and some will be shorter.

After you have judged the reference line, assign a number to each following line so that it represents how long the line is in proportion to the reference. The longer it is compared to the reference, the larger the number you will use; the shorter it is compared to the reference, the smaller the number you will use. So if you feel that a line is twice as long as the reference, give it a number twice the reference number; if it’s a third as long, provide a number a third as big as the reference.

So if the reference is this line, you might give it the number 10:

If you have to judge this line, you might assign it 17:

And this one might be 2.5:

There is no limit to the range of numbers you may use. You may use whole numbers or decimals. If you assigned the reference line the number 1, you might want to c last one 0.25. Just try to make each number match the length of the line as you see it.

Parts 2 and 3: Judging Sentences

In Part 1 of the experiment you used numbers to estimate the length of lines on the screen. In parts 2 and 3 you will use numbers to judge the ease of comprehension of some English sentences in the same way.
You will see a series of sentences presented one at a time on the screen. Each sentence is different. Some will seem perfectly okay to you, but others will not. Your task is to judge how easy or difficult it is to comprehend each sentence by assigning a number to it.

As with the lines in Part 1, you will first see a reference sentence, and you can use any number that seems appropriate to you for this reference. For each sentence after the reference, you will assign a number to show how easy or difficult that sentence is in proportion to the reference sentence, which will also appear with your rating.

For example, if the reference sentence was:

(1) The dog devoured yesterday the bone.

you would give it a certain number of your choosing. If the next example:

(2) The ate the bone dog.

seemed 10 times more difficult to comprehend than the reference, you’d give it a number 10 times the number you gave to the reference. If it seemed half as good as the reference, you’d give it a number half the number you gave to the reference.

You can use any range of positive numbers that you like, including decimal numbers. You can use any number, except that you cannot use zero or negative numbers. Try to use a wide range of numbers and to distinguish as many degrees of acceptability as possible.

There are no ‘correct’ answers, so whatever seems right to you is a valid response. We are interested in your first impressions, so please don’t take too much time to think about any one sentence: try to make up your mind quickly, spending less than 10 seconds on a sentence.

In each part you will see the reference item in the experiment window. Please enter your reference number and then press the ENTER key. Now the test items will appear one after the other in the experiment window. Please type your judgment in the box below for each item.

The experiment will take 10 to 15 minutes.

Please keep in mind:
* Use any number you like for the reference sentence.
* Judge each sentence in proportion to the reference.
* Use any positive numbers, which you think are appropriate.
* Use high numbers for ‘difficult’ sentences, low numbers for ‘easy’ sentences and intermediate numbers for sentences, which are intermediate in their ease or difficulty of comprehension.
* Try to use a wide range of numbers and to distinguish as many degrees of acceptability as possible.
* Try to make up your mind quickly, base your judgments on your first impressions.
Appendix IV

List of test sentences for experiment 2

The first sentence in each pair is the theta-role surplus version; the second is its disambiguated version.

1. a. Zoo visitors threw the gorillas which were eating some crunchy and tasty peanuts.
   b. Zoo visitors observed the gorillas which were eating some crunchy and tasty peanuts.

2. a. The posh lady promised the gentleman who was dancing the lively and elegant Waltz.
   b. The posh lady liked the gentleman who was dancing the lively and elegant Waltz.

3. a. Nurses serve old patients who can readily swallow some solid and chewy food.
   b. Nurses mind old patients who can readily swallow some solid and chewy food.

4. a. The friendly doctor gave the child who was chewing the colorful and sweet toffee.
   b. The friendly doctor examined the child who was chewing the colorful and sweet toffee.

5. a. A field reporter told the anchorman who was reading the horrible and outrageous news.
   b. A field reporter briefed the anchorman who was reading the horrible and outrageous news.

6. a. The parent got the boisterous child who was playing the entertaining and captivating game.
   b. The parent ignored the boisterous child who was playing the entertaining and captivating game.

7. a. The bidder sold the poor woman who was losing the ugly and fake diamond.
   b. The bidder ignored the poor woman who was losing the ugly and fake diamond.

8. a. Some volunteers bought the penniless man who was moving some old and used furniture.
   b. Some volunteers helped the penniless man who was moving some old and used furniture.

9. a. The generous host passed the guest-chef who was cooking the smelly and slippery fish.
   b. The generous host assisted the guest-chef who was cooking the smelly and slippery fish.
10. a. A young waitress poured the visitors who were already drinking some red and fruity wine.
   b. A young waitress attended the visitors who were already drinking some red and fruity wine.

11. a. The young boy brought the handyman who was changing the round and white light bulb.
    b. The young boy watched the handyman who was changing the round and white light bulb.

12. a. The PhD student offered his colleague who was writing the long and complex paper.
    b. The PhD student supported his colleague who was writing the long and complex paper.

The first sentence in every pair is the locally ambiguous version; the second is its disambiguated version.

1. a. While the ferocious baron was hunting the frightened deer ran into the woods.
    b. While the ferocious baron was sleeping the frightened deer ran into the woods.

2. a. While the reporter was photographing the flamboyant celebrity posed on the podium.
    b. While the reporter was resting the flamboyant celebrity posed on the podium.

3. a. While the sophomore student was reading the glossy magazine blew off the desk.
    b. While the sophomore student was yawning the glossy magazine blew off the desk.

4. a. While the policemen were investigating the drug dealers tried to hide the stolen narcotics.
    b. While the policemen were coming the drug dealers tried to hide the stolen narcotics.

5. a. While the secretary was typing the long letter arrived in the mail with the postman.
    b. While the secretary was talking the long letter arrived in the mail with the postman.

6. a. While the pupil was studying the important article was posted on the internet.
    b. While the pupil was idling the important article was posted on the internet.
7. a. While the princess was paying the helpful footman opened the carriage door for her.
b. While the princess was awakening the helpful footman opened the carriage door for her.

8. a. While my friend Rick was driving the big truck veered into a ditch on the side of the road.
b. While my friend Rick was chatting the big truck veered into a ditch on the side of the road.

9. a. After the tourist drank the tap water proved to be contaminated with poisonous chemicals.
b. After the tourist departed the tap water proved to be contaminated with poisonous chemicals.

10. a. While the new chef was stirring the onion soup boiled over in the long silver pot.
b. While the new chef was snoozing the onion soup boiled over in the long silver pot.

11. a. While the known professor was lecturing the new students took notes on their laptops.
b. While the known professor was speaking the new students took notes on their laptops.

12. a. While the modern artist was painting the beautiful model sat on the sofa at the studio.
b. While the modern artist was thinking the beautiful model sat on the sofa at the studio.
Appendix V

List of test sentences for experiment 3, part A

The first sentence in each triplet is the locally ambiguous version, the second is its disambiguated version, and the third is the theta surplus version.

1. a. De voetganger attendeerde de automobiliste die vanwege haar signalen had genegeerd op het kind.
   b. De voetganger attendeerde de automobiliste die vanwege haar de signalen had genegeerd op het kind.
   c. De voetganger attendeerde de automobiliste die al enkele minuten had gewacht op het kind.

2. a. De werker vertelde de schoonmaakster die met haar gordijnen had versteld over de familieproblemen.
   b. De werker vertelde de schoonmaakster die met haar de gordijnen had versteld over de familieproblemen.
   c. De werker vertelde de schoonmaakster die af en toe had geroddeld over de familieproblemen.

3. a. De politieagent associeerde de zwerfster die van haar dekens had ontvangen met een dealer.
   b. De politieagent associeerde de zwerfster die van haar de dekens had ontvangen met een dealer.
   c. De politieagent associeerde de zwerfster die gisteravond nog hevig had gevochten met een dealer.

4. a. De docent verwees de studente die van haar dossiers had geleend naar de studieadviseur.
   b. De docent verwees de studente die van haar de dossiers had geleend naar de studieadviseur.
   c. De docent verwees de studente die altijd vol aandacht had geluisterd naar de studieadviseur.

5. a. De galeriehouder wees de kunstenares die aan haar schilderijen had verkocht op een standbeeld.
   b. De galeriehouder wees de kunstenares die aan haar de schilderijen had verkocht op een standbeeld.
   c. De galeriehouder wees de kunstenares die zojuist zeer fors had afgedongen op een standbeeld.

6. a. De adviseur herinnerde de advocate die met haar memo's had doorgelezen aan het jaarverslag.
   b. De adviseur herinnerde de advocate die met haar de memo's had doorgelezen aan het jaarverslag.
   c. De adviseur herinnerde de advocate die in de avonden had gewerkt aan het jaarverslag.
7. a. De leraar scheidde de leerlinge die met haar examens had uitgedeeld van een buurman.
b. De leraar scheidde de leerlinge die met haar de examens had uitgedeeld van een buurman.
c. De leraar scheidde de leerlinge die altijd alle examens had overgeschreven van een buurman.

8. a. De ontwerper behing de misdienaars die voor haar sieraden had gedragen met de kettingen.
b. De ontwerper behing de misswinnares die voor haar de sieraden had gedragen met de kettingen.
c. De ontwerper behing de misswinnares die overdag zonder doel had rondgelopen met de kettingen.

9. a. De directeur verzocht de secretaresse die met haar ideeën had uitgewisseld om het rapport.
b. De directeur verzocht de secretaresse die met haar de ideeën had uitgewisseld om het rapport.
c. De directeur verzocht de secretaresse die de hele bijeenkomst had gelachen om het rapport.

10. a. De bejaarde man stuurde de dochter die voor haar brieven had opgeruimd naar de apotheker.
b. De bejaarde man stuurde de dochter die voor haar de brieven had opgeruimd naar de apotheker.
c. De bejaarde man stuurde de dochter die nooit een brief had verzonden naar de apotheker.

11. a. De verkoper voorzag de muziekliefhebster die van haar kadobonnen had gekocht van de cd.
b. De verkoper voorzag de muziekliefhebster die van haar de kadobonnen had gekocht van de cd.
c. De verkoper voorzag de muziekliefhebster die al heel wat had gehoord van de cd.

12. a. De begeleider vertelde de wandelaarster die met haar schoenen had geruild over de looproutes.
b. De begeleider vertelde de wandelaarster die met haar de schoenen had geruild over de looproutes.
c. De begeleider vertelde de wandelaarster die al heel wat had gelezen over de looproutes.
Appendix VI
Plausibility and acceptability pilot experiment, Experiment 3

In the plausibility section, the test sentences were broken up into three groups in the following manner. The theta-role surplus condition was divided into two conditions. The High Verb Attachment (HVA) tested whether the high attachment of the final PP was plausible. The Low Verb Attachment (LVA) tested whether the final PP could be plausibly attached to the lower verb. The Object Drop (OD) condition tested whether the omission of the object was possible to make sure the ambitransitive verb could drop its prepositional argument and still maintain plausibility.

(201) **High Verb Attachment (HVA)**

De begeleider vertelde de wandelaarster over de looproutes.
The guide (Mas) told the hiker (Fem) about the trail.

‘The guide told the hiker about the trail.’

(202) **Low Verb Attachment (LVA)**

De wandelaarster had heel wat gelezen over de looproutes.
The hiker (Fem) had quite a lot read about the trail.

‘The hiker had read quite a lot about the trail.’

(203) **Object Drop (OD)**

De wandelaarster had heel wat gelezen.
The hiker (Fem) had quite a lot read.

‘The hiker had read quite a lot.’

The rating of plausibility was between 1 and 5 (1 = implausible, 5 = plausible). The experiment design was based on (3 conditions) x (14 items) = 42 sentences. Those were randomized into four lists, containing 55 fillers each of the same sort of structure but having various other verbs.

In the acceptability test, the theta-role surplus condition was broken up into two. The High Attachment (HA) condition was intended to affirm the acceptability of the final PP attachment to the upper verb (204). The Object Drop – High Verb (ODHV) (205) tested whether the absence of an object would result in unacceptability of the upper ditransitive verb.

To exclude the possible attachment of the final PP to the lower verb in the locally ambiguous condition, which might augment the already present haart ambiguity in the locally ambiguous condition, the Low Verb Illicit Attachment (LVIA) condition was constructed (206), which was supposed to result in unacceptability. In Dutch, it is possible to attach the final PP to a preceding NP. Special care was taken in the lexical choice of the PP to disallow such option. The final PP was omitted from the ambitransitive verb to see whether object absence still resulted in acceptability (207).
(204) \textit{High Attachment (HA)}

De begeleider vertelde de wandelaarster over de looptroutes.
The guide (Mas) told the hiker (Fem) about the trail.

‘The guide told the hiker about the trail.’

(205) \textit{Object Drop – High Verb (ODHV)}

De begeleider vertelde de wandelaarster.
The guide (Mas) told the hiker (Fem).

‘The guide told the hiker.’

(206) \textit{Low Verb Illicit Attachment (LVIA)}

De wandelaarster had schoenen geruild over de looptroutes.
The hiker (Fem) had shoes swapped about the trail.

‘The hiker had swapped shoes about the trail.’

(207) \textit{Object Drop (OD)}

De wandelaarster had heel wat gelezen.
The hiker (Fem) had quite a lot read.

‘The hiker had read quite a lot.’

The rating of acceptability was between 1 and 5 (1 = unacceptable, 5 = acceptable).
The experiment model was based on (4 conditions) x (14 items) = 56 sentences.
Those were randomized into 4 lists, containing 56 fillers each, manipulated in the same fashion, only with various other verbs. In total, the questionnaire contained 98 experimental sentences randomized among 111 fillers.

The order in half of the questionnaires was first the plausibility test then the acceptability test, in the other half, the order was reversed. The instructions to the participants for both the plausibility and acceptability tests are found in Appendix VII.

Data from 85 participants, different from the ones mentioned in Section §2.7.2, were collected. All were students at Utrecht University, who voluntarily participated in the experiment. Their average age was 21; all were native speakers of Dutch.

The average plausibility of the OD condition was 4.5, MSe = 0.4 (collapsed over items). It can be assumed then that OD items were plausible. No significant difference in plausibility was found between conditions HVA (Mean = 4.4, MSe = 0.05) and LVA (Mean = 4.3, MSe = 0.05): $t_1 (84) = 1.236, p < 0.220, MSe = 0.036$; $t_2 (13) = 0.553, p < 0.590, MSe = 0.083$. The conclusion is that these conditions were equally plausible.

A significant difference in acceptability was found between conditions HA (Mean = 4.7, MSe = 0.04) and ODHV (Mean = 2.6, MSe = 0.07) conditions: $t_1 (84) = 26.808, p < 0.001, MSe = 0.079$; $t_2 (13) = 9.919, p < 0.001, MSe = 0.213$. Therefore HA is acceptable, ODHV is much less so: the upper verb requires a PP
argument. The illicit low attachment of the final PP in the locally ambiguous condition resulted in unacceptability as the low mean score of LVIA indicates (Mean = 2.6, MSe = 0.07). Condition OD was acceptable (Mean = 4.7, MSe = 0.04). The conclusion is that the final PP in the theta-role surplus condition can be selected by the lower and upper verbs, and that the lower verb may omit the final PP. Furthermore, the attachment of the final PP to the lower verb in the locally ambiguous condition is neither plausible nor acceptable, such that it does not augment difficulty. This ensures then that considerations of plausibility and acceptability of attachments will not act as complicating factors in attachment decisions.
Appendix VII

Instructions to participants, Plausibility and Acceptability (Dutch)

The following pages contain a list of sentences, divided into two sections. The first section tests the plausibility of sentences; the second section tests their correctness.

Read each sentence as you would normally do to understand what it means. Then evaluate the sentence on a scale of (1) to (5) by circling the chosen number.

In the plausibility section, give a high number (such as 5) when the sentence describes a situation that is likely to occur. Give a low number (such as 1) when the sentence describes a situation that is unlikely to occur. Make as much use of all scale values!

A sample of an unlikely or unlikely sentence is (a):

(a) The teacher separated the student from the floor.

It is easy to see that you would evaluate sentence (a) with a low number (1), because separating a student from the floor is unlikely. The following sentence:

(b) The teacher separated the student from a fellow student.

correctly describes a situation that is likely to occur. You should, therefore, evaluate this sentence with a high number (such as 5).

In the correctness section, give a high number (such as 5) when the sentence is correct, that is, grammatically correct (written in good Dutch). Give a low number (such as 1) when the sentence is incorrect. Again, make use of all scale values!

A sample of an incorrect sentence is (c):

(c) The teacher separated the student.

Sentence (c) is incorrect, the sentence should describe what the teacher would separate the student. You could therefore evaluate sentence (c) for example with a (2). Sentence (a), although not plausible, is correct, and it could therefore be evaluated in the correctness section. It describes the situation that the teacher would separate the student, namely the floor.

It is important that the number you give to a sentence is an indication of your first impression. Do not spend too much time on the sentences. Read a sentence quickly, and immediately give your evaluation. Then move on to the next sentence. Work through the list at a brisk pace.

Thank you for your cooperation!
Appendix VIII

Instructions to subjects in experiment 3, parts A and B (Dutch)

Je staat op het punt mee te doen aan een psycholinguïstisch onderzoek dat test hoe makkelijk zinnen te begrijpen zijn, ofwel hoe moeilijk jij het vindt om bepaalde zinnen te begrijpen.

Het onderzoek is alleen bedoeld voor moedertaalsprekers van het Nederlands, en duurt ongeveer 20 á 25 minuten. Alvast bedankt voor het meedoen.

Het onderzoek bestaat uit drie onderdelen.

In het eerste onderdeel zal je als oefening de lengte van lijnen beoordelen. Ten eerste verschijnt er een referentielijn. Je dient de lengte van die lijn in te schatten door er een getal naar eigen keuze aan toe te kennen.

Als de lijn bijvoorbeeld 6 centimeter lang lijkt, kun je er het getal 6 aan geven. Hierna zal er een serie lijnen met verschillende lengtes verschijnen. Je dient de lengte van deze lijnen in te schatten in vergelijking met de referentielijn. Als de nieuwe lijn bijvoorbeeld drie keer zo lang als de referentielijn lijkt, kun je er het getal 18 aan geven. Als de lijn twee keer zo kort lijkt geef je het getal 3 eraan.

In het tweede onderdeel zullen jezelf trainen in het beoordelen van de complexiteit van een aantal oefenzinnen in vergelijking met een referentiezin. Om een oordeel te geven over de complexiteit van een zin moet je kijken of de zin durend de zin begrijpen, en of je de zin moest herlezen.

Als eerste zal er een referentiezin verschijnen, en een vraag over deze zin. Je dient deze vraag te beantwoorden door op het juiste antwoord te klikken. Je krijgt hierna direct te zien of je het goede antwoord hebt gegeven of niet. Het is niet mogelijk om je antwoord aan te passen nadat je het antwoord hebt gecorrigeerd. Hierna dien je de complexiteit van de referentiezin te beoordelen door er een getal naar eigen keuze aan toe te kennen.

Vervolgens zullen er nog enkele zinnen verschijnen. Over elk van deze zinnen dien je een vraag te beantwoorden, en de complexiteit ervan te beoordelen in vergelijking met de referentiezin door er een getal aan toe te kennen, net als bij de lijnen uit het eerste deel.

Het derde en laatste onderdeel is het onderzoeksdeel. In dit onderdeel zullen jij de zelfde taak uitvoeren als in onderdeel twee, alleen dan met testzinnen.

Houd in gedachten dat het lijkt alsof sommige zinnen fout of vreemd zijn, of ten onrechte geen interpunctie hebben. Dit is opzettelijk gedaan, het maakt deel uit van het experiment!

Let op:

* Beoordeel elke zin IN VERGELIJKING met de referentiezin;
* Gebruik ELK WILLEKEURIG POSITIEF GETAL (MAAR GEEN NUL) voor de zinnen;
* Gebruik HOGE getallen voor ‘complexe’ zinnen, LAGE getallen voor ‘niet-complexe’ zinnen, en getallen ERTUSSENIN voor zinnen die tussen ‘complex’ en ‘niet-complex’ in zitten;
* Gebruik een GROOT BEREIK van getallen;
* Vorm snel je mening, besteed NOOIT MEER DAN 10 SECONDEN aan één zin, en baseer je oordeel op je eerste indruk.
Appendix IX

Detailed instructions (Dutch)


Je wordt gevraagd om de volgende gegevens op te geven:
* Je naam en e-mail adres;
* Je leeftijd en geslacht;
* Of je een linguïst bent of niet;
* Je moedertaal.

Het experiment begint nadat je op de ‘BEGINNEN’ knop hebt geklikt. Het experiment bestaat uit de volgende drie onderdelen:
1. Trainingsonderdeel: beoordelen van lijnlengtes
2. Oefenonderdeel: beoordelen van zinnen
3. Onderzoeksonderdeel: beoordelen van zinnen

Onderdeel 1: beoordelen van lijnlengtes

Voordat je aan het hoofdonderdeel van het onderzoek begint, zul je een korte taak uitvoeren die betrekking heeft op het beoordelen van lijnlengtes. Een reeks lijnen van verschillende lengtes zal op het scherm weergegeven worden. Het is je taak om te bepalen hoe lang ze zijn door ze een getal toe te kennen. Het is daarbij de bedoeling dat je schattingen in vergelijking worden gemaakt met de eerste lijn die je zult zien, de zogenaamde referentielijn. Je dient daarom de referentielijn een willekeurig getal toe te kennen. Houd hierbij wel in gedachten dat sommige lijnen langer, en andere lijnen korter zullen zijn dan de referentielijn.

Nadat je de referentielijn beoordeeld hebt, dien je een getal toe te kennen aan de daarop volgende lijnen. Dit getal moet weergeven hoe lang de lijn is in vergelijking met de referentielijn. Hoe langer deze lijn is in vergelijking met de referentielijn, des te hoger dient het getal te zijn dat je toekent. Hoe korter de lijn is in vergelijking met de referentielijn, des te lager dient het getal te zijn dat je toekent. Als je bijvoorbeeld vindt dat een lijn tweemaal zo lang is als de referentielijn, dan geef je een getal dat tweemaal zo hoog is als het getal dat je aan de referentielijn hebt toegekend. Als de lijn driemaal zo kort is, dan geef je een getal dat drie maal zo laag is als de referentielengte.

Als dit bijvoorbeeld de referentielijn is, dan kun je deze het getal 10 geven:

Als je deze lijn moet beoordelen, dan kun je deze het getal 17 geven:

En deze lijn kan dan 2,5 zijn:
Er zijn geen grenzen aan het bereik van getallen dat je mag gebruiken. Je mag zowel hele getallen als decimalen geven. Als je in bovenstaand voorbeeld de referentielijn het getal 1 hebt toegekend, dan kun je de laatste lijn het getal 0,25 geven. Het enige dat echt belangrijk is, is dat elk getal dat je geeft overeenkomt met de lengte van de lijn zoals jij deze ziet.

Onderdeel 2 en 3: beoordelen van zinnen

In onderdeel 1 van het onderzoek heb je getallen gebruikt om de lengte in te schatten van lijnen op je scherm. In onderdeel 2 en 3 zul je op dezelfde manier getallen gebruiken om uit te drukken hoe makkelijk of moeilijk het begrip van een aantal Nederlandse zinnen zijn. Ook zal je gevraagd worden om begripsvragen over de zinnen te beantwoorden. Je zult een reeks zinnen zien die één voor één weergegeven worden. Elke zin is verschillend. Sommige zullen volledig correct op je overkomen, maar anderen niet. Het is je taak om te beoordelen hoe makkelijk of moeilijk het is om deze zinnen te begrijpen. Dit doe je door elke zin een getal toe te kennen.


Je mag elk getal gebruiken dat je geschikt lijkt voor deze referentiezin. Na elke zin na de referentiezin dien je een begripsvraag te beantwoorden, en dien je een getal toe te kennen dat uitdrukt hoe makkelijk of moeilijk deze zin is om te begrijpen in vergelijking met de referentiezin. Deze referentiezin en jouw toegewezen getal zullen ook verschijnen.

Als bijvoorbeeld dit je referentiezin is:

(1) De hond gisteren verslond het bot.

Dan zou je deze zin een getal toekennen. Als het volgende voorbeeld:

(2) De at het bot hond

je 10 keer zo moeilijk zou lijken als de referentiezin, dan zou je deze zin een getal geven dat 10 maal het getal is dat je aan de referentiezin hebt gegeven. Als je de zin half zo moeilijk vindt als de referentiezin, dan zou je een getal geven dat de helft is van de referentiezin.

Een voorbeeld van een begripsvraag zou kunnen zijn:

(3) Verslond de hond het bot?


Er zijn geen grenzen aan het bereik van getallen dat je mag gebruiken. Je mag zowel hele getallen als decimalen geven. De enige beperking is dat je geen negatieve getallen of het getal 0 mag geven. Probeer een zo groot mogelijk bereik van getallen te gebruiken, en probeer zo veel mogelijk verschillende gradaties van acceptabelheid te onderscheiden.

Er zijn geen ‘juiste’ oordeelen, dus je beoordeling is geldig indien het uitdruk wat jij vindt. We zijn geïnteresseerd in je eerste indruk. Neem daarom niet te veel tijd om over één zin na te denken. Vorm zo snel mogelijk je mening, en besteed niet meer dan 10 seconden aan één zin.

In elk van de drie onderdelen zal je als eerste het referentieobject in het onderzoeksscherm zien. Geef je referentiegetal op, en druk dan op de ENTER toets. Hierna zullen de test-items
in het onderzoeksscherm verschijnen. Geef je beoordelingsgetal in het kader daaronder voor elk item. Vergeet niet de begripsvraag te beantwoorden die na elke zin verschijnt.

Dit experiment duurt 10 á 15 minuten.

Houd tijdens het onderzoek in gedachten:

* Je mag elk getal geven aan de referentielijn en de referentiezin;
* Je dient elke zin te beoordelen in verhouding tot de referentiezin;
* Je mag elk positiief getal gebruiken dat je geschikt lijkt;
* Je dient hoge getallen te geven aan ‘moeilijke’ zinnen, lage getallen aan ‘makkelijke’ zinnen, en gemiddelde getallen aan zinnen die gemiddeld zijn qua moeilijkheidsgraad bij het begrijpen ervan;
* Je dient een groot bereik van getallen te gebruiken;
* Je dient zo snel mogelijk je mening te vormen, en je oordelen te baseren op je eerste indruk.
Appendix X

List of test sentences for experiment 3, part B

The first sentence in each triplet is the locally ambiguous version, the second is the theta surplus version, and the third is its disambiguated version.

1. a. De voetganger attendeerde de automobiliste die vanwege haar signalen had genegeerd op het kind.
   b. De voetganger attendeerde de automobiliste die al enkele minuten had gewacht op het kind.
   c. De voetganger attendeerde de automobiliste, die al enkele minuten had gewacht, op het kind.

2. a. De werker vertelde de schoonmaakster die met haar gordijnen had versteld over de familieproblemen.
   b. De werker vertelde de schoonmaakster die af en toe had geroddeld over de familieproblemen.
   c. De werker vertelde de schoonmaakster, die af en toe had geroddeld, over de familieproblemen.

3. a. De politieagent associeerde de zwerfster die van haar dekens had ontvangen met een dealer.
   b. De politieagent associeerde de zwerfster die gisteravond nog hevig had gevochten met een dealer.
   c. De politieagent associeerde de zwerfster, die gisteravond nog hevig had gevochten, met een dealer.

4. a. De docent verwees de studente die van haar dossiers had geleend naar de studieadviseur.
   b. De docent verwees de studente die altijd vol aandacht had geluisterd naar de studieadviseur.
   c. De docent verwees de studente, die altijd vol aandacht had geluisterd, naar de studieadviseur.

5. a. De galeriehouder wees de kunstenares die aan haar schilderijen had verkocht op een standbeeld.
   b. De galeriehouder wees de kunstenares die zojuist zeer fors had afgedongen op een standbeeld.
   c. De galeriehouder wees de kunstenares, die zojuist zeer fors had afgedongen, op een standbeeld.

6. a. De adviseur herinnerde de advocate die met haar memo's had doorgelezen aan het jaarverslag.
   b. De adviseur herinnerde de advocate die in de avonden had gewerkt aan het jaarverslag.
   c. De adviseur herinnerde de advocate, die in de avonden had gewerkt, aan het jaarverslag.
7. a. De leraar scheidde de leerlinge die met haar examens had uitgedeeld van een buurman.
b. De leraar scheidde de leerlinge die altijd alle examens had overgeschreven van een buurman.
c. De leraar scheidde de leerlinge, die altijd alle examens had overgeschreven, van een buurman.

8. a. De ontwerper behing de misswinnares die voor haar sieraden had gedragen met de kettingen.
b. De ontwerper behing de misswinnares die overdag zonder doel had rondgelopen met de kettingen.
c. De ontwerper behing de misswinnares, die overdag zonder doel had rondgelopen, met de kettingen.

9. a. De directeur verzocht de secretaresse die met haar ideeën had uitgewisseld om het rapport.
b. De directeur verzocht de secretaresse die de hele bijeenkomst had gelachen om het rapport.
c. De directeur verzocht de secretaresse, die de hele bijeenkomst had gelachen, om het rapport.

10. a. De bejaarde man stuurde de dochter die voor haar brieven had opgeruimd naar de apotheker.
b. De bejaarde man stuurde de dochter die nooit een brief had verzonden naar de apotheker.
c. De bejaarde man stuurde de dochter, die nooit een brief had verzonden, naar de apotheker.

11. a. De verkoper voorzag de muziekliefhebster die van haar kadobonnen had gekocht van de cd.
b. De verkoper voorzag de muziekliefhebster die al heel wat had gehoord van de cd.
c. De verkoper voorzag de muziekliefhebster, die al heel wat had gehoord, van de cd.

12. a. De begeleider vertelde de wandelaarster die met haar schoenen had geruild over de looproutes.
b. De begeleider vertelde de wandelaarster die al heel wat had gelezen over de looproutes.
c. De begeleider vertelde de wandelaarster, die al heel wat had gelezen, over de looproutes.
Appendix XI

List of sentences for experiment 4

In each quintet, the first sentence belongs to condition 2SE2OR, the second 1SE2OR, the third 2SE1OR, the fourth 2SE, and the last is Sequential.

1. a. The warden that the prisoner that the social worker visited hated deprived food of prisoners.
b. The warden forwarded the demand that the prisoner that the social worker visited made.
c. The report that the prisoners that the social worker visited ran away was alarming to many citizens.
d. The disclosure that the report that the prisoners ran away was true alarmed many citizens.
e. The social worker visited the prisoners that hated the warden that didn’t give them enough food.

2. a. The mother that the child that the teenager looked after depended on took care of children.
b. The mother took care of the child that the teenager that the agency recommended looked after.
c. The news that the child that the teenager looked after died yesterday moved the community.
d. The confirmation that the message that the child died was true shocked the community.
e. The teenager looked after the child that depended on the mother that gave orphans many Christmas presents.

3. a. The singer that the pianist that the fans admired toured with had a beautiful voice.
b. The fans supported the singer that the pianist that the critics recommended toured with.
c. The review that the singer that the fans admired coughed damaged her reputation.
d. The review that the story that the singer coughed was published was damaging.
e. The fans admired the pianist that toured with the singer that had a beautiful voice.
4. a. The nurse that the hospital supervisor that the doctor reported to fired lost her work permit too.
b. The doctor trained the nurse that the hospital supervisor that the committee-head consulted fired.
c. The rumor that the nurse that the doctor trained lost her work permit astonished many people.
d. The rumor that the report that the nurse lost her work permit was true astonished many people.
e. The doctor informed the hospital supervisor that fired the nurse that lost her work permit too.

5. a. The actor that the producer that the reviewer applauded supported had a successful career.
b. The producer admired the actor that the reviewer that the theatre sent praised.
c. The article that the actor that the reviewer admired played well delighted the producer.
d. The news that the article that the actor played well was supportive delighted the producer.
e. The reviewer applauded the producer that supported the actor that had a failing career.

6. a. The nanny that the agent that the colleague recommended sent had a lot of experience.
b. The agent recommended the maid that the colleague that the secretary called hired.
c. The fact that the nanny that the colleague recommended had a lot of experience was already known.
d. The message that the rumor that the maid cleans well was true was already known.
e. The colleague recommended the agent that sent the nanny that had a lot of experience.

7. a. The boy that the teacher that the principal cheered up disciplined committed a crime again.
b. The principal cheered up the teacher that the boy that the inspector examined disappointed.
c. The rumor that the boy that the teacher disciplined was a criminal reached the students.
d. The discovery that the report that the boy committed a crime was true alarmed many parents.
e. The principal praised the teacher that disciplined the boy that committed a crime again.
8.  
a. The clergyman that the bishop that the journalist denounced appointed molested altar-boys regularly.
b. The bishop denounced the clergyman that the journalist that the altar-boy called accused.
c. The news that the clergyman that the journalist deplored was a molester was appalling.
d. The news that the story that the clergyman molested altar-boys was true appalled the bishop.
e. The journalist deplored the bishop that appointed the clergyman that molested children.

9.  
a. The defendant that the lawyer that the district-attorney appointed cross-examined incriminated himself in the trial.
b. The defendant consulted the lawyer that the district attorney that the judge warned appointed.
c. The fact that the defendant that the district-attorney cross-examined incriminated himself was worrying.
d. The fact that the report that the defendant incriminated himself was true was posted everywhere.
e. The district attorney appointed the lawyer that cross-examined the defendant that incriminated himself several times during the trial.

10.  
a. The actress that the director that the studio manager hired admired won an award for her role.
b. The director admired the actress that the critic that the editor hired disparaged.
c. The story that the actress that the director hired won an award did not surprise the fans.
d. The announcement that the news that the actress won an award was true was not surprising.
e. The studio manager hired the director that admired the actress that won an award for playing a homeless person.
Appendix XII

Instructions to subjects in experiments 4, 5

You are about to participate in a psycholinguistic experiment that tests the ease of comprehension of sentences, that is how difficult you find sentences to understand. The experiment is intended only for native speakers of English, and takes about 15-20 minutes.

If you are only curious what it is all about and do not intend to finish the experiment, please do not start the experiment at all. This is simply because unfinished attempts mess up the experiment design.

The experiment consists of three parts. In the first part, you will rate the length of lines, as practice. First a reference line will appear, and you will have to assign it any number of your choosing. Additional lines will appear, which you will have to judge in proportion to that reference line.

For example, you decided to rate the reference line with 4, corresponding to the line’s length in centimeters or inches. If the following line seems twice as big, you can give it an 8.

In the second part, you will train on rating the ease of comprehension of a few practice sentences in comparison to a reference sentence. The principle is the same as in the first part. For example, you decided that the difficulty of the reference sentence was 10. If the following, practice sentence is 2.5 times more difficult than the reference sentence, give the practice sentence a higher rating, maybe 25.

The third and last part is the experimental part and there you will have to rate test sentences in proportion to a reference sentence as you did in the second part.

NOTE: Some of the sentences might appear to be grammatically incorrect, missing commas, or totally wrong. That’s part of the experiment. Try to do your best and trust your intuitive judgment!

Please keep in mind:
* To judge each sentence IN PROPORTION to the reference.
* To use ANY POSITIVE NUMBER (and NO ZEROS) you like for the reference sentence/line and the following sentences/lines.
* To use HIGH numbers for ‘difficult’ sentences, LOW numbers for ‘easy’ sentences and INTERMEDIATE numbers for sentences, which are between ‘difficult’ and ‘easy’.
* To try to use a WIDE RANGE of numbers and to distinguish as many degrees of ease of comprehension as possible.
* To try to make up your mind QUICKLY, never spend more than 10 seconds on a sentence, and base your judgments on your first impressions.
Appendix XIII
List of sentences for experiment 5
High Interference is the first of each pair, Low Interference the second.

1. a. The announcement that the rumor that the claim was shocking was true was exciting.
   b. The announcement that the rumor that the actress won an award was true was exciting.

2. a. The fact that the claim that the idea was scary was true was disappointing.
   b. The fact that the claim that the guy eats garlic was true was disappointing.

3. a. The rumor that the report that the story was accurate was false was deplorable.
   b. The rumor that the report that the immigrant lost his work permit was false was deplorable.

4. a. The discovery that the report that the precedent was important was true was alarming.
   b. The discovery that the report that the boy committed a crime was true was alarming.

5. a. The confirmation that the report that the rumor was malicious was false was surprising.
   b. The confirmation that the report that the president fired a general was false was surprising.

6. a. The fact that the rumor that the insinuation was accurate was malicious was well-known.
   b. The fact that the rumor that the babysitter stole money was malicious was well-known.

7. a. The belief that the fact that the idea was valid was well-founded was strong.
   b. The belief that the fact that the singer committed suicide was well-founded was strong.

8. a. The fact that the rumor that the belief was false was true was important.
   b. The fact that the rumor that the boxer lost a round was true was important.

9. a. The expectation that the belief that the suspicion was correct was reliable was high.
   b. The expectation that the belief that the prince had an affair was reliable was high.
10. a. The story that the fact that the idea was innovative was true was inspiring.
b. The story that the fact that the painter had many ideas was true was inspiring.

11. a. The report that the suggestion that the fact was accurate was misleading was contradictory.
b. The report that the suggestion that the landlord evaded taxes was misleading was contradictory.

12. a. The conclusion that the report that the complaint was legitimate was correct was official.
b. The conclusion that the report that the patient had cancer was correct was official.

13. a. The insinuation that the belief that the idea was bad was wrong was offensive.
b. The insinuation that the belief that the casino changed the cards was wrong was offensive.

14. a. The report that the conclusion that the principle was accurate was rash was shocking.
b. The report that the conclusion that the researcher found a skeleton was rash was shocking.

15. a. The idea that the agreement that the belief was false was common was shocking.
b. The idea that the agreement that the soldiers evacuated squatters was common was shocking.

In each triplet, the first sentence is Stage, Short Version, the second Event, the third Stage, Long Version.

1. a. The claim that the rumor that the mayor quit circulated surprised voters.
b. The claim that the rumor that the mayor was a gangster circulated surprised voters.
c. The claim that the rumor that the mayor unexpectedly quit circulated surprised voters.

2. a. The message that the claim that the alien appeared spread interested us.
b. The message that the claim that the alien was green was spread interested us.
c. The message that the claim that the alien suddenly appeared spread interested us.

3. a. The fact that the report that the lawyer lied appeared disturbed the public.
b. The fact that the report that the lawyer was a liar appeared disturbed the public.
c. The fact that the report that the lawyer knowingly lied appeared disturbed the public.
4. a. The fact that the suggestion that the macho-guy cheated appeared damaged the girl.
b. The fact that the suggestion that the macho-guy was a cheater appeared damaged the girl.
c. The fact that the suggestion that the macho-guy intentionally cheated appeared damaged the girl.

5. a. The confirmation that the message that the prisoner escaped appeared worried the community.
b. The confirmation that the message that the prisoner was a con-artist appeared worried the community.
c. The confirmation that the message that the prisoner unexpectedly escaped appeared worried the community.

6. a. The belief that the rumor that the child choked emerged concerned many parents.
b. The belief that the rumor that the child was helpless emerged concerned many parents.
c. The belief that the rumor that the child accidentally choked emerged concerned many parents.

7. a. The belief that the dream that the pilot crashed persisted bothered the villagers.
b. The belief that the dream that the pilot was reckless persisted bothered the villagers.
c. The belief that the dream that the pilot knowingly crashed persisted bothered the villagers.

8. a. The complaint that the insinuation that the leader hesitated spread materialized.
b. The complaint that the insinuation that the leader was hesitant spread materialized.
c. The complaint that the insinuation that the leader repeatedly hesitated spread materialized.

9. a. The claim that the rumor that the guard snoozed spread reached the headlines.
b. The claim that the rumor that the guard was lazy spread reached the headlines.
c. The claim that the rumor that the guard regularly snoozed spread reached the headlines.

10. a. The belief that the suspicion that the lady sinned materialized grew.
b. The belief that the suspicion that the lady was a sinner materialized grew.
c. The belief that the suspicion that the lady often sinned materialized grew.

11. a. The fact that the news that the baby drowned broke upset parents.
b. The fact that the news that the baby was helpless broke upset parents.
c. The fact that the news that the baby suddenly drowned broke upset parents.
12. a. The rumor that the belief that the cub survived spread delighted naturalists. 
   b. The rumor that the belief that the cub was a survivor spread delighted naturalists. 
   c. The rumor that the belief that the cub surprisingly survived spread delighted naturalists. 

13. a. The fact that the claim that the lecturer lied circulated amazed the students. 
   b. The fact that the claim that the lecturer was a liar circulated amazed the students. 
   c. The fact that the claim that the lecturer continually lied circulated amazed the students. 

14. a. The news that the suggestion that the mammoth existed prevailed interested many people. 
   b. The news that the assumption that the mammoth was huge prevailed interested many people. 
   c. The news that the suggestion that the mammoth actually existed prevailed interested many people. 

15. a. The fact that the rumor that the ghost vanished existed remained a mystery. 
   b. The fact that the rumor that the ghost was scary existed remained a mystery. 
   c. The fact that the rumor that the ghost abruptly vanished existed remained a mystery.
References


Samenvatting

Het belangrijkste doel van deze dissertatie is aan te tonen dat problemen met taalverwerking die optreden bij zogenaamde garden path-zinnen en constructies met center-embedding te herleiden zijn op de lokale schending van een grammaticaal principe, het Theta Principle. Vervolgens wordt betoogd dat dit Theta Principle psychologische realiteit heeft.

In deze dissertatie worden drie stellingen beargumenteerd. Ten eerste, thematische overwegingen zijn van invloed op beslissingen over de eerste analyse (initial attachment) door het menselijke zinsverwerkingsysteem (de parser). Ten tweede hebben de verwerkingsproblemen die voorkomen bij center-embedding een syntactische verklaring, in tegenstelling tot eerdere analyses uit het generatieve onderzoek van de laatste vijftig jaar. De ontdekking dat center-embedded zinnen opgevat kunnen worden als sterke eilanden leidt ertoe dat de verwerkingsproblemen verklaard kunnen worden door het interventie-effect dat optreedt bij eilandconstructies. Dit interventie-effect zorgt ervoor dat het lokaal onmogelijk is om aan het Theta Principle te voldoen, waardoor men moeilijkheden ondervindt bij de verwerking van dit soort zinnen. Ten derde kunnen de verwerkingsproblemen die optreden bij beide typen zinnen nu verenigd worden binnen één theoretische verklaring, aangezien ze beide voortkomen uit lokale schendingen van een globaal grammaticaal principe, het Theta Principle.

Om deze argumenten te staven wordt gebruik gemaakt van een thematische parser - het strong transparent semi head-driven model - gebaseerd op Pritchett (1992a). Dit model is (sterk) transparant omdat parsing performance direct afgeleid wordt uit grammatical competence. Het model is semi head-driven omdat aanhechting wordt beïnvloed door informatie over het hoofd; de parser kan evenwel wel hypothesen opstellen over de inputstring die overeenkomen met zowel de informatie over het hoofd als de grammatica. Het belangrijkste principe waardoor de parser zich laat leiden bij aanhechtingen is Theta Attachment; de parser probeert zo snel mogelijk de thematische structuur van de input vast te stellen.

In Hoofdstuk 1 worden argumenten weergegeven die tegen dit type model van menselijke taalverwerking zijn ingebracht, en die voornamelijk stammen uit OV-talen. Ter aanvulling op Prichett’s originele theorie worden de verwerkingsproblemen die optreden bij garden path constructies verklaard door een lokale schending van een globaal grammaticaal principe, het Theta Principle.

In Hoofdstuk 2 wordt de ‘goodness of fit’ van het model aangetoond met behulp van experimentele studies. Het model doet enkele specifieke voorspellingen met betrekking tot het optreden van het garden path-effect in een bijzonder zinstype, de zogenaamde theta-role surplus zinnen. Een voorbeeld van dit type zinnen is *A waitress poured the visitors who were drinking some wine*. Het zinsdeel *some wine* kan zowel laag (aan *drink*) of hoog (aan *pour*) worden aangehecht. Als de parser besluit tot lage aanhechting (low attachment) wordt een garden path effect voorspeld. Als er echter voor hoge aanhechting wordt gekozen (high attachment) zal
er een grammaticale zin worden gevormd waardoor er geen heranalyse hoeft plaats te vinden. Om deze reden zal hoge aanhechting dan ook niet leiden tot een garden path-effect. Aangezien Theta Attachment niet kan beslissen tussen hoge of lage aanhechting, is de voorspelling dat lezers inconsequent zullen zijn bij het rapporteren van een garden path effect, dat immers afhankelijk is van een willekeurige beslissing van de parser. Dit betekent dan ook dat het garden path-effect willekeurig zal optreden. Deze voorspelling kan als volgt worden uitgelegd: het gedrag van mensen bij het lezen van theta-role surplus zinnen zal verschillen van standaard garden path-zinnen (zoals After Fred eats food gets thrown) en (gemakkelijke) zinnen zonder garden path.

De voorspellingen van andere verwerkingsmodellen met betrekking tot theta-role surplus zinnen worden vervolgens gepresenteerd, zoals het Garden Path Model (Frazier, 1987b), het Diagnosis Model (Fodor and Inoue, 1994), Structural Determinism (Gorrell, 1995a, Sturt and Crocker, 1996), het LAST Model (Townsend and Bever, 2001), het Snip Model (Lewis, 1993), en het Memory-Based Ranked Parallel Model (Gibson, 1991). Al deze modellen doen voorspellingen die afwijken van de thematische parser, wat theta-role surplus zinnen geschikt maakt om dit model van de andere modellen te onderscheiden.

Om de voorspellingen te toetsen was het nodig om een geschikte experimentele methode te vinden en deze ook te valideren. Aangezien het garden path-effect te maken heeft met een lokaal oordeel van ongrammaticaliteit is voor Magnitude Estimation gekozen, een techniek die subtiele gradaties in grammaticaliteitsoordelen kan meten. Het voordeel van deze methode is dat het complexe parametrische statistische analyses mogelijk maakt, iets wat niet mogelijk is wanneer vragenlijsten met meerkeuzevragen gebruikt worden.

Omdat het om een off-line methode gaat, is eerst bekeken of de techniek ook gebruikt kan worden voor meten van real-time effecten zoals het garden path effect. In een experiment werd proefpersonen gevraagd zinnen te beoordelen op basis van complexiteit. Zinnen uit een oogbewegingsstudie in het Engels (Frazier & Rayner, 1982) werden gebruikt om de resultaten van de twee experimentele methoden te vergelijken. Hieruit bleek dat Magnitude Estimation de resultaten van het oogbewegingsexperiment replicateerde. Bovendien bleek het een gevoelige methode om zelfs milde verwerkingsproblemen en lengte-effecten te meten. Concluderend kan gesteld worden dat de methode voldeed om parsing performance te onderzoeken.

In een tweede Magnitude Estimation experiment werd gevonden dat het garden path effect willekeurig optrad in Nederlandse theta-role surplus zinnen (De begeleider vertelde de wandelaarster die al heel wat had gelezen over de looproutes), en dat zulke zinnen significant verschillenden van standaard garden path zinnen en zinnen zonder garden path. De gemiddelde moeilijkheidsgraad van theta role surplus-zinnen lag precies tussen die van de beide andere typen zinnen in.

In dit hoofdstuk is dan ook empirisch bewijs gevonden voor de veronderstelling dat de thematische parser overeenkomt met het werkelijke verwerkingsgedrag van mensen op theta-role surplus zinnen, terwijl andere modellen deze data niet kunnen verklaren. In dit hoofdstuk wordt geconstateerd dat
Theta Attachment een leidend principe van syntactische aanhechting is, en dat zinsverwerking gedreven wordt door thematische overwegingen.

In Hoofdstuk 3 wordt de mogelijkheid onderzocht of het model een groter empirische bereik heeft dan alleen het voorspellen van garden path effecten, door ook center-embedding te onderzoeken. Canonieke gevallen van center-embedding behelsjen dubbele inbeddingen van relatieve bijzinnen waarin het relativum als object fungeert, zoals The man the cat the dog chased bit died. De syntactische verklaring van verwerkingsproblemen die optreden bij deze canonieke gevallen steunt op de observatie dat ze beschouwd kunnen worden als een complex NP-eiland ingebed in een complex NP-eiland. Het was dan ook eerst nodig om te laten zien dat deze syntactische eigenschap relevant is voor de verwerkingsproblemen. Het was daarnaast ook nodig om andere hypotheses over de oorzaak van deze verwerkingsproblemen te weerleggen.

Twee Magnitude Estimation-experimenten werden speciaal ontwikkeld om verschillende hypotheses van andere modellen met betrekking tot center-embedding te toetsen. Het eerste experiment was bedoeld om te onderzoeken of de eilandstatus van canonieke center-embedding-zinnen ertoe leidt dat ze als significant moeilijker beoordeeld worden. Als voorbeelden van canonieke gevallen met eilandstatus werd gekozen voor dubbele center-embedded relatieve bijzinnen die als een adjunct aan het gerelativiseerd nomen aangehecht zijn. De resultaten lieten zien dat dit soort zinnen als moeilijker werden beoordeeld dan zinnen met dubbel ingebedde bijzinnen die als een complement aan het nomen aangehecht zijn, zoals the disclosure [that the report [that the prisoner ran away] was true] alarmed many citizens. Bovendien sloot het experiment de mogelijkheid uit dat het hoge aantal ingebedde zinnen al dan niet in combinatie met object-relatieve bijzinnen de oorzaak van de moeilijkheid waren. Het tweede experiment toonde aan dat de voorspellingen van de Dependency Locality Theory (Gibson, 2000) en de Similarity-Based Theory (Lewis, 1996) niet overeenkwamen met de resultaten. Er werden geen significante verschillen gevonden als de aan- of afwezigheid van discourse referenten werd gemanipuleerd (als in de Dependency Locality Theory) of als de overeenkomst tussen de initiële NP’s werd bekeken. Vervolgens werden deze theorieën verder uitgesloten door te laten zien dat ze ook de verkeerde voorspellingen doen met betrekking tot andere constructies.

Verwerkingsproblemen die optreden bij center-embedding zijn een uiting van een lokale schending van het Theta Principle. De thematische operator van de meest ingebedde relatieve bijzin (the cat Op, the dog chased ei) blokkeert de operator van de hogere bijzin (the man Op, the cat... bit ei), die de gap-positie niet kan binden (ei). Dit is een van de belangrijkste bevindingen van deze dissertatie - een interventie effect in canonieke gevallen van center-embedding dat lijk op het effect dat wordt aangetroffen in andere complexe NP-eilanden. De interventie blokkeert de mogelijkheid om aan het Theta Principle te voldoen, waardoor verwerkingsproblemen ontstaan. Op deze gronden wordt betoogd dat zinnen met center-embedding lokaal ongrammaticaal zijn, en niet onverwerkbaar (unprocessable). Dit is een nieuwe claim, omdat zinnen met center-embedding traditioneel gezien worden als onverwerkbaar, maar wel globaal grammaticaal.
Het model kan ook verklaren dat canonieke *center-embedding* ‘beter’ wordt door eerste en tweede persoons voornaamwoorden te gebruiken. Ook kan het model het ontbreken van verwerkingsproblemen bij dubbele inbeddingen van vrije relatieve en dubbele inbeddingen van subjectzinnen verklaren. Er vindt eenvoudigweg geen interventie plaats tijdens het verwerken van deze constructies. Het betrekkelijke gemak waarmee center-embedding in links vertakkende talen wordt verwerkt, wordt ook verklaard door de *thematic parser*.

Hoofdstuk 4 behelst een beknopte samenvatting van de eerdere hoofdstukken, en verbindt het garden path effect met het effect van *center-embedding*. De *thematic parser* verklaart het feit dat garden path zinnen relatief gemakkelijk zijn ten opzichte van de moeilijke canonieke gevallen van *center-embedding*. Deze laatsten veroorzaken de zwaarste verwerkingsproblemen omdat door interventie niet aan het *Theta Principle* kan worden voldaan. Interventie treedt echter niet op bij garden path constructies, waardoor de verwerking minder problematisch is.

In navolging van Pritchett (1992a) wordt betoogd dat een *semi head-driven* benadering van zinsverwerking die direct geformuleerd is in termen van grammaticale principes, cognitief voldoet. Het model dekt een breed scala aan data uit zowel rechts- als linksvertakkende talen, en op basis hiervan wordt betoogd dat het menselijk natuurlijke taalverwerkingsmechanisme universeel is. De parser kan geheel in termen van grammatica beschouwd worden, terwijl grammatica niet als een systeem van regels gezien wordt maar als een aantal beperkingen op representaties, en de parser de grammaticale principes lokaal toepast op elke punt gedurende het verwerken van de inputstring.

Het is van cruciaal belang dat er een overeenkomst is vastgesteld tussen het niet voldoen aan het Theta Principle en de externe werkelijkheid – een fenomeen dat fysieke realiteit heeft in het menselijk brein – uitgedrukt in verwerkingsproblemen bij garden path zinnen en *center-embedding*. Dit effect is meetbaar en ‘werkelijk’; het *Theta Principle* heeft psychologische realiteit en is een werkelijke eigenschap van de menselijke geest.
Curriculum Vitae

Oren Sadeh-Leicht was born on January 7, 1973 in Petah-Tikvah, Israel. He matriculated in honors from the ‘Gimnasia Realit’ High School in Rishon-Letsion in 1991 and started his studies at Tel-Aviv University.

He obtained his B. Sc. degree in Chemistry in 1994. Between 1995 and 1999 he served as an Academic Professional Officer in the Israeli Procurement Corps.

After completing his military service, he worked as a freelance translator while studying for his MA degree in Linguistics at Tel-Aviv University. He received his MA degree in 2003. At the same year, he started his employment as a graduate student at the Utrecht Institute of Linguistics OTS in the Netherlands. This dissertation is the result of the work he carried out there.