ATTACHMENT AMBIGUITIES IN HEBREW COMPLEX NOMINALS:

PROSODY AND PARSING

by

AMIT SHAKED

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2009

© 2009

AMIT SHAKED

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Linguistics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Janet Dean Fodor

Date — Chair of Examining Committee

Gita Martohardjono

Date — Executive Officer

Dianne C. Bradley

Marcel den Dikken

Eva M. Fernández

Tal Siloni

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

ATTACHMENT AMBIGUITIES IN HEBREW COMPLEX NOMINALS:

PROSODY AND PARSING

by

Amit Shaked

Adviser: Professor Janet Dean Fodor

This dissertation investigates the prosody-syntax interface in two Hebrew complex nominals. Their meaning is identical but their syntax and phonology differ. The *free state* (FS) nominal is similar to English (e.g., *the coach of the wrestler*). The *construct state* (CS) nominal lacks the preposition, has stress on the second noun, and the construction constitutes one phonological word despite its internal syntactic structure. When followed by a relative clause (RC) the two constructions (1a,b) are syntactically/semantically ambiguous in the same way; but they contrast prosodically because one of the phrasing patterns permitted for FS (a boundary between the two nouns) is inhibited for CS.

(1)	a. FS:	ha-me'amen	šel	ha-mit'agref	še-paraš le'axar ha-taxarut
	b. CS:	me'amen		ha-mit'agref	še-paraš le'axar ha-taxarut
		(the-)coach	of	the-wrestler	who-retired after the-fight

Research on the relative clause attachment ambiguity has shown that different languages and constructions vary in attachment preferences; it is important to investigate factors that may contribute to attachment decisions, because this variation threatens the universality and innateness of the human sentence processing routines. In recent years, much research has focused on a prosodic explanation: The Implicit Prosody Hypothesis (Fodor, 2002) proposes that attachment choices in silent reading are affected by implicit prosody mentally projected onto the sentence by readers, on the basis of language-specific phonological rules. Since implicit prosody cannot be directly observed, it is investigated by studying overt prosody in speaking and listening. This dissertation thus provides data on prosodic phrasing in Hebrew and its relation to the processing of the RC-attachment construction.

Three experiments are reported: a production experiment studied speakers' preferred prosodic patterns for each construction; a listening study examined how the preferred interpretation differed for different prosodic contours; a silent reading experiment investigated the interpretation when no overt prosody was provided. For the FS sentences the results of the three experiments concurred, indicating that RC-attachment is sensitive to prosody, both explicit and implicit. For the CS, prosody was also found to play a role but some other factor(s) intrinsic to the construction resulted in consistently higher RC-attachment.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my dissertation adviser, Dr. Janet Dean Fodor, for her ceaseless guidance and support from the conception to the completion of this dissertation. I am indebted to her for her enlightening questions, brilliant thinking and for making sense of the intricate data collected for this project. I feel honored that she shared with me her time, her insight and her expertise.

I am also grateful to members of my committee, Dr. Dianne Bradley, Dr. Eva Fernández and Dr. Marcel den Dikken, without whom the ideas pursued here would not have fully developed. Especially, I would like to express my great appreciation to Dianne for providing a great deal of helpful advice, constructive criticism, and important suggestions throughout the development of this work, and for always having an open door. In addition, I would like to thank Dr. Tal Siloni, who served as an external examiner on this dissertation, for her interest in this project and for sharing her vast knowledge on the syntax of Hebrew complex nominals.

I am also grateful to other members of the faculty at the Linguistics program of the CUNY Graduate Center, in particular Dr. Charles Cairns and Dr. Gita Martohardjono, who served as Executive officers during my graduate-school life. I thank Chuck for opening a window when a door was closed, and Gita for always offering an open ear and wise advice, and for showing me how much fun academic life can be. Each was a vital source of support and encouragement, which I desperately needed and greatly appreciate.

I wish to express my gratitude to to the faculty at the department of linguistics at Tel Aviv University in Israel for enabling me to recruit and test students in their facility,

in particular Dr. Julia Horvath and Dr. Tal Siloni, who initially welcomed me into the department and Dr. Mira Ariel and Dr. Rachel Giora for generously volunteering their offices for testing participants in subsequent experiments. I am also grateful to students at Tel-Aviv University who participated in the experiments reported in this dissertation.

The completion of the experiments would not have been possible without the help of several people in various stages of this research. I would like to thank the trained speaker in my listening experiment, Efrat Shani Levin, for volunteering her time and effort, and for her endless patience in ensuring that the recording satisfied all criteria. Many thanks to Dafna Laadan and Efrat Friedmann Shaked, who were willing to spend a large amount of time recruiting participants and administering the listening experiment in Israel. My deep appreciation also goes to Keren Ohayon, for her time and assistance in the initial segmentation of the production data and in providing the ear judgments for the production data as well as the listening stimuli. I would also like to thank Ronit Webman for the numerous conversations on Hebrew prosody and linguistics.

To my parents, I owe my deepest gratitude. Thank you for your unconditional, unwavering and selfless love and support throughout my life, which have endowed me with the courage and resolve to confront all challenges in my life.

Finally, I am endlessly grateful to my husband for his love and support, for giving me strength and making me a better person, and to our amazing son, Noam, who is the greatest accomplishment in my life.

TABLE OF CONTENTS

ABSTRACT…… iv

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS viii

LIST OF TABLES xii

LIST OF FIGURES xiv

CHAPTER 1 INTRODUCTION AND OVERVIEW 1

CHAPTER 2 HEBREW COMPLEX NOMINALS: AN OVERVIEW 6

2.1 Semantic relations and stylistic variations 7

2.2 Syntactic properties of FS and CS: Similarities and differences 12

2.3 The phonological properties of the CS 19

2.4 Hebrew compounds 26

2.5 The syntax of Hebrew complex nominals 40
2.5.1 The Free State 44
2.5.2 The Construct State 49

2.6 Summary ……56

CHAPTER 3 HEBREW PROSODY: AN OVERVIEW 58

3.1 Inventory of nuclear tones and intonation contours in Hebrew 58

3.2 Phrase boundaries in Hebrew 65

3.3 Stress assignment in the Hebrew nominal system 72

3.4 Summary.. 77

CHAPTER 4 MATERIALS AND GENERAL DESIGN OF EXPERIMENTS 79

4.1 Materials. 81

4.2 Rephrasability assessment 85
4.2.1 Design 86
4.2.2 Procedure 87
4.2.3 Participants 89
4.2.4 Item selection procedure 90

4.3 RC attachment assessment 91
4.3.1 Design 91
4.3.4 Procedure 93
4.3.3 Results 95

Appendix A to Chapter 4: Rephrasability assessment 99

Appendix B to Chapter 4: RC attachment assessment 110

CHAPTER 5 PRODUCTION OF DISAMBIGUATED SENTENCES 115

5.1 Aims…… 115

5.2 Method…. 122
5.2.1 Materials 122
5.2.2 Design 124
5.2.3 Procedure 127
5.2.4 Participants 130
5.2.5 Data measurement and assessment 131
5.2.5.1 Perceived phrasing data 131
5.2.5.2 Acoustic data 138

5.3 Analysis and results 144
5.3.1 Perceived phrasing data 144
5.3.2 Acoustic data 155

5.4 Summary and conclusions 165

Appendix A to Chapter 5: Variety of phrasing patterns observed in the production study and their assignment to three phrasing categories. 169

Appendix B to Chapter 5: Incidence of phrasing categories established for the analysis of the production study. 170

Appendix C to Chapter 5: Perceived breaks over complete utterances. 174

CHAPTER 6 PERCEPTION OF SPOKEN SENTENCES 177

6.1 Goals…… 177

6.2 Method…. 184
6.2.1 Materials and design 184
6.2.2 Recording and analysis of auditory stimuli 194
6.2.3 Procedure 199
6.2.4 Participants 201
6.2.5 Data measurement and assessment 202

6.3 Analysis and results 203

6.4 Conclusions 209

Appendix A to Chapter 6: Consideration of position for additional prosodic breaks prior to the complex nominal with RC. 214

Appendix B to Chapter 6: Considerations in the choice of RC length 219

Appendix C to Chapter 6: Instrumental analysis of recorded stimuli 223

CHAPTER 7 ATTACHMENT IN SILENT READING 229

7.1 Aims…… 229

7.2 Method…. 234
7.2.1 Materials and design 234
7.2.2 Procedure 238
7.2.3 Participants 241
7.2.4 Data measurement and assessment 241

7.3 Analysis and results 242

7.4 Summary of attachment findings 244

7.5 Integrating the outcomes of the three experiments for the FS 245

7.6 Understanding the data for the CS 253

CHAPTER 8 CONCLUSIONS 258

8.1 Motivation and aims 260

8.2 Summary of experimental findings 264
8.2.1 The Free State 264
8.2.2 The Construct State 267
8.2.3 Major conclusions based on experimental findings 269

8.3 Addressing the CS puzzle 270

8.4 Future directions 290

APPENDICES…. 297

Appendix A: Materials 297
Appendix A-1: Target items used in the three experiments for this dissertation 297
Appendix A-2: Target items in the production experiment 303
Appendix A-3: Target items in the listening experiment 304
Appendix A-4: Target items in the silent reading experiment 305
Appendix A-5: Filler items used in the experiments for this dissertation 306

Appendix B: Instructions to participants 325
Appendix B-1: Instructions for the production experiment 325
Appendix B-2: Instructions for the listening experiment 327
Appendix B-3: Instructions for the silent reading experiment 329

Appendix C: Language history (background) questionnaire 331

Appendix D: Summary tables for Analyses of Variance 333
Appendix D-1: Production experiment 333
Appendix D-2: Listening experiment 354
Appendix D-3: Silent reading experiment 359

REFERENCES…… 361

LIST OF TABLES

Table 3-1: Inventory of Hebrew nuclear tones, as presented in Laufer (1987) in Hebrew (pp.32-33) and translated into English by AS. The symbols are to be read from right to left (following the directionality in Hebrew orthography).....59

Table 3-2: Sentence intonation contours in Hebrew, translated from Laufer (1987), with a short description of each contour. 62

Table 4-1: Percentage of plausibility/acceptability of N1 and N2 in CS and in FS. 96

Table 5-1: Proposed break indices for Hebrew and their definitions, modeled on the ToBI system of Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005). 134

Table 5-2: A notation system for identifying Hebrew intermediate phrase boundaries. Definitions are based on (Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005), with my adaptations. 136

Table 5-3: Segmentation of waveform into 6 regions comparable between the 4 experimental conditions. 140

Table 6-1: Experimental conditions tested in the listening experiment, as a function of prosodic phrasing (Phrasing 0/ Phrasing 1/ Phrasing 2), construction type (CS/FS), and RC length (short/long RC). A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3).....185

Table 6-2: Comparison of FS utterances produced with three phrasing patterns: Phrasing 0, Phrasing 1 and Phrasing 2. RC contains 1 prosodic word uniformly.....187

Table 6-3: A fully crossed comparison of construction type (FS or CS) and phrasing patterns (Phrasing 0, Phrasing 1 and Phrasing 2), with RC uniformly 1 prosodic word.....188

Table 6-4: A comparison of RC length (short and long) for CS utterances produced with Phrasing 2. A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3). 188

Table 6-5: Factorial design of the listening experiment varying phrasing pattern, construction type and RC length, with conditions omitted on feasibility/grammaticality grounds shaded in dark gray, and a condition omitted on economy grounds shaded in light gray. A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3). 189

Table 7-1: Summary of results for FS production, listening and silent reading experiments. The measures here mix prosody rates (in the production study) and attachment rates in the listening and silent reading studies. 246

Table 7-2: summary of results for CS production, listening and silent reading experiments. The measures here mix prosody rates (in the production study) and attachment rates in the listening and silent reading studies. 254

LIST OF FIGURES

Figure 2-1: FS, underlying structure, from Siloni (1996, 1997).....47

Figure 2-2: FS, surface structure from Siloni (1996, 1997).....47

Figure 2-3: CS, underlying structure, from Siloni (1996, 1997) 52

Figure 2-4: CS, surface structure, from Siloni (1996, 1997) 52

Figure 5-2: Derived structures of FS, with high and low RC associations.....118

Figure 5-3: Derived structures of CS, with high and low RC associations. 119

Figure 5-4: Waveform and f_0 contour for the Hebrew utterance *ha-ˈohadim heˈericu ˈet ha-meˈamen šel ha-mitˈagref še-paraš* 'The fans admired the coach of the wrestler who retired'. The left box indicates the N1 region, and the right box indicates the N2 region. 142

Figure 5-5: Percentages of perceived breaks for high attachment utterances, categorized by prosodic break position, as a function of construction type and RC length.....147

Figure 5-6: Percentage of perceived breaks for low attachment utterances, categorized by prosodic break position, as a function of construction type and RC length.....147

Figure 5-7: Percentage of perceived phrasing for FS utterances, categorized by prosodic break position, as a function of attachment height, and RC length.....150

Figure 5-8: Percentage of perceived phrasing for CS utterances, categorized by prosodic break position, as a function of attachment height, and RC length.....153

Figure 5-9: Valley-to-peak f_0 rise for short and long RC, as a function of construction type, RC length and attachment height, for N1.....156

Figure 5-10: Valley-to-peak f_0 rise for short and long RC, as a function of construction type, RC length and attachment height, for N2.....156

Figure 5-11: Duration of the N1 region, as a function of construction type, RC length and attachment height. 157

Figure 5-12: Duration of the N2 region, as a function of construction type, RC length and attachment height. 157

Figure 6-1: Mean duration for the N1 region in each of the 6 experimental conditions tested in the listening study. 196

Figure 6-2: Mean duration for the N2 region in each of the 6 experimental conditions tested in the listening study. 197

Figure 6-3: Design 1, FS Listening data in terms of % high attachment, for each of the 3 phrasing patterns: Phrasing 0, Phrasing 1 and Phrasing 2. RC contains 1 prosodic word, uniformly. 204

Figure 6-4: Design 2, Listening data in terms of % high attachment, for CS and FS in two phrasing patterns: Prosody 0 and Prosody 2. RC length is fixed at 1 prosodic word. 205

Figure 6-5: Listening data in terms of % high attachment, for CS sentences with Phrasing 2 in two RC lengths: Short RC (1 prosodic word) and long RC (3 prosodic words). 207

Figure 7-1: RC attachment to FS and CS in silent reading in Hebrew 242

Figure 8-1: Restrictive RC association in the CS 286

Figure 8-2: Restrictive RC association in the FS 287

CHAPTER 1

INTRODUCTION AND OVERVIEW

This research involves the ambiguity resolution of Relative Clause attachment to complex nominals in Hebrew. The Hebrew language has a variety of complex nominal constructions, thus lending itself to a within-language investigation of the prosodic properties related to this construction and their relation to its attachment interpretation. In particular, this research focuses on two complex nominals: the Free State nominal (FS) and the Construct State nominal (CS).[1] The FS construction has a preposition *šel* equivalent to the English *of* between the two nouns, and is generally assumed to have the same syntactic structure as possessive constructions in English. The CS construction, in comparison, has no such element within the complex nominal, and has several other unique phonological and syntactic properties that differentiate it from the FS, as will be discussed below. Examples of the FS and the CS are presented in (1) and (2), respectively.

(1) ha-me'amen šel ha-mit'agref
the-coach of the-wrestler

(2) me'amen ha-mit'agref
coach the-wrestler

[1] Another type of possessive construction is available in Hebrew: the double genitive nominal. This construction has two possessive markers: an element *šel* ('of') between the nouns and a genitive marker on N1, agreeing with N2 in gender and number. An example is presented in (i). This double genitive nominal will be mentioned briefly in chapter 2 but will not be investigated in this dissertation (see Chapter 2).

(i) me'amno šel ha-mit'agref
coach(gen) of the-wrestler
'The coach of the wrestler'

Importantly, while the two constructions have distinct properties, their meanings are essentially identical. Both the CS and FS in the examples above carry the meaning: 'the coach of the wrestler'.

The purpose of this dissertation is to examine Hebrew speakers' understanding of sentences containing these two constructions when followed by a relative clause (RC) that ambiguously modifies one or the other of the nouns in the complex nominal, as illustrated in (3)[2].

(3) (ha-)me'amen (šel) ha-mit'agref še-paraš le'axar ha-taxarut
(the-)coach (of) the-wrestler who-retired after the-fight
'The coach of the wrestler who retired after the fight'

Sentences containing a complex nominal followed by an RC as in (3) are in principle ambiguous between two readings, i.e. either the whole complex nominal headed by the first noun (*me'amen* 'coach') or the second noun on its own (*mit'agref* 'wrestler') can be understood as modified by the RC.

The puzzle presented by the relative clause attachment ambiguity has been a focus of psycholinguistic research for more than 20 years because speakers of different languages have been shown to exhibit different ambiguity resolution preferences (Cuetos & Mitchell, 1988). This cross-linguistic variation is of theoretical importance because it threatens the claim that the human sentence processing routines are universal and innate. It is essential, therefore, to assess factors that have been considered in models of sentence processing, in order to explain why the above variation occurs and whether the empirical

[2] The term 'complex nominal' is used throughout this dissertation to refer to the noun sequences that constitute the FS and CS, even though in fact the RC is part of the syntactic structure of these complex nominals. When discussing the sequence of N1-N2-RC I will refer to the RC specifically, as in 'a complex nominal with an RC'.

data can still be reconciled with the innateness of the human sentence processing mechanism. Hebrew can contribute to this exploration, because it offers two complex nominals which have the same meaning and yet differ phonologically and syntactically. The within language contrast in the syntactic and phonological factors can therefore be distinguished experimentally from semantic biases in ambiguity resolution.

In the history of the field of psycholinguistics, no other construction has been tested across languages more than the RC construction. Several explanations have been offered to account for the cross-linguistic variation in RC attachment preferences, including an explanation based on Construal Theory (Frazier & Clifton, 1996), an Anaphoric Binding dual-mechanism explanation (Hemforth & colleagues, e.g., Hemforth et al., 1998), an explanation based on the Tuning Hypothesis (Cuetos, Mitchell & colleagues, e.g., Mitchell & Cuetos, 1991; Cuetos et al., 1996), and a parameter setting explanation based on Recency/ Predicate proximity (Gibson & colleagues, e.g., Gibson et al., 1996). In recent years, a large part of the discussion has focused on prosodic accounts for the explanation of the cross-linguistic variations. Lovrić (2003) and Augurzky (2005) present arguments showing that the Implicit Prosody Hypothesis (IPH) proposed by Fodor (2002) can explain the data for Croatian and German, respectively, but other approaches to attachment ambiguity resolution cannot. For reasons of space, I focus on the IPH and do not review the alternative proposed explanations here; they have been thoroughly discussed and disputed in recent work by Lovrić (2003) and Augurzky (2005).

The IPH (Fodor, 2002) claims that in silent reading implicit prosody is computed and may affect attachment preferences of ambiguous sentences. The syntactic

configuration associated with the default prosodic contour for a given construction will be favored by the parser. While these and other studies convincingly demonstrate the effect of prosody on attachment interpretation, they do not always present a detailed description of the prosodic facts related to this construction, a crucial stage in relating the silent reading facts to prosodic phrasing. Thus, there is room for a more thorough investigation of prosody itself, and its role in the interpretation of this construction. This dissertation investigates the prosodic interface and its relation to the interpretation of the RC construction in Hebrew.

Chapter 2 of the dissertation provides a thorough review of the literature on the two complex nominals under investigation. First, I address the semantic relations expressed by the two constructions and stylistic differences in the use of the FS and CS nominals. I then describe the characteristics of each complex nominal and point out the similarities and differences between them. Then, I discuss in particular the unique phonological properties of the CS, which bear on the analysis to follow. Discussion of the CS cannot be complete without the mention of Hebrew nominal compounds, which are closely related to the CS. I present the similarities and differences between CS and compounds, and list the various approaches in the literature to their analysis. Finally, I present a syntactic analysis of the two complex nominals. The Hebrew CS has been a focus of syntactic research for many years, and there are several competing analyses available. I focus on the syntactic analysis proposed by Siloni (1996, 1997), while mentioning only briefly alternative analyses.

Chapter 3 of this dissertation reviews previous work on Hebrew prosody. While Hebrew prosody at the sentential level has not been a focus of study in the past, I will

attempt to provide a detailed picture of the main issues concerning the phonology of Hebrew prosody, which serve as the basis for the prosodic analysis in the current investigation. I present an inventory of the nuclear tones and intonation contours in Hebrew and the characteristics associated with each; I discuss cues for identifying phrase boundaries; and present an overview of stress assignment in the Hebrew nominal system, which is relevant for the discussion of the status of the CS as a single phonological word.

Chapter 4 introduces the experimental component of this dissertation. It describes and motivates the general design of the materials used for the set of experiments. Chapter 5 presents data on the production of disambiguated RC sentences. It assesses the prosodic properties related to the Hebrew RC construction and provides a base for subsequent discussion of the possible effect of prosody on the interpretation of ambiguous sentences. It also considers the characteristics of Hebrew prosody in general. Chapter 6 and Chapter 7 present data on the interpretation of ambiguous RC sentences: Chapter 6 presents listening data on attachment preferences of spoken sentences, while Chapter 7 provides data on the attachment preferences in silent reading. Together, the experimental chapters provide data which can be used to assess contributing factors to the RC attachment interpretation in Hebrew. The evaluation of the results from the three experiments will be discussed in Chapter 7. The conclusion to this dissertation (Chapter 8) summarizes the main findings of the dissertation and addresses the interesting puzzle presented by the CS data.

CHAPTER 2

HEBREW COMPLEX NOMINALS: AN OVERVIEW

Hebrew has a number of complex nominals expressing a similar meaning but varying in their structural representation: The Free State nominal (FS) has a preposition *šel* equivalent to the English 'of' between the two nouns, while each of the nouns heads its own maximal projection. The Construct State nominal (CS) does not have the preposition *šel* between the two nouns, and the first noun in the sequence is in a reduced (construct) form – it is dependent on the second noun for phonological stress and definiteness marking, as will be discussed in detail below (§2.2). There is another complex nominal in Hebrew, the Double Genitive construction, which is mentioned here briefly but is not under investigation in this dissertation, since this construction is infrequent in Hebrew (Berman, 1988; Ravid et al., 1986).[3] It includes a pronominal suffix (e.g., *–o* 'his') on the first noun in addition to a preposition *šel* between the two nouns. The three complex nominals are presented in (1).

[3] In a corpus study, Ravid & Shlesinger (1995) found that double genitives are very rare both in oral (3% of all genitives) and written (7.5% of all genitives) discourse, even in highly formal registers of Hebrew (compared with the occurrence of the CS and the FS in the oral (58% and 39%, respectively) and written corpora (78% and 15%, respectively).

(1) a. *Free State*

ha-me'amen šel ha-mit'agref
the-coach of the-wrestler

b. *Construct State*

me'amen ha-mit'agref
coach the-wrestler

c. *Double Genitive*

me'amno šel ha-mit'agref
coach.his$_i$ of the-wrestler$_i$

Although their syntactic representations vary, the three complex nominals have the same meaning. The meaning conveyed in all the complex nominals in (1) is 'the coach of the wrestler' (in English, equivalently 'the wrestler's coach'). This chapter provides a review of the literature on the semantic, phonological and syntactic aspects of these complex nominals, specifically the FS and CS, which are the focus of investigation in this dissertation. Readers who are familiar with the properties of the Hebrew FS and CS constructions may prefer to skip this survey of the literature. For readers interested primarily in the experimental data, the brief summary in Chapter 1 of points essential to the design of the experiments may suffice.

2.1 Semantic relations and stylistic variations

It is largely agreed in the contemporary literature on Hebrew complex nominals that the three alternate complex nominal forms are essentially synonymous and can express the same range of semantic relations (Anderson, 1985; Azar, 1976; Berman, 1978, 1988; Glinert, 1989; Cohen, 1992; Ritter, 1991, among others). The earlier

literature on the Hebrew genitive constructions attempted to establish structural criteria for the preference of one complex nominal over the other (see, for instance, Rosen 1957, 1966, Ornan 1964, and more recent discussion in Doron 1989).[4] Berman (1978), on the other hand, explicitly denies any systematic semantic difference between the two complex nominals; on the basis of reports of native informants, she claims that the three forms are interchangeable and that the choice between them is stylistic in nature, rather than a semantic difference, as discussed further below (see also Levi, 1976 for similar claims, and more recently Siloni, 1997).

In terms of semantic relations conveyed by the genitive constructions, much of the literature has focused on a detailed classification of the semantic relations conveyed by the CS. Avineri (1964), Reif (1968), Levi (1976) and Glinert (1989), among others, provide a list of various relations that may be expressed using the CS, such as material (2a), possessive (2b), action/state (2c), purpose (2d), content (2e), measure/container (2f) and causal relations (2g), among others. The examples in (2) can all be expressed using an FS as well. Compare, for example, with comparable facts for Spanish and English in Gilboy et al. (1995).

(2) a. kos plastik
glass plastic
'A plastic cup'

b. mexonit ha-sar
car the-minister
'The minister's car'

[4] One attempt to associate different semantic relations with each of the complex nominals is the claim made by Rosen (1957, 1966) that FS nominals are preferred when N2 is [+HUMAN] and N1 is not (expressing inalienable relations), similarly to the Saxon genitive construction in English, in which the possessor or adjunct is generally human (e.g., *the man's car*), as opposed to the use of *of* elsewhere (e.g., *the door of the car,* rather than *the car's door*).

c. tiftuf ha-berez
dripping the-tap
'The dripping of the tap'

d. simlat xatuna
dress wedding
'A wedding dress'

e. argaz xol
box sand
'A sandbox'

f. karton xalav
carton milk
'A milk carton'

g. laxac avir
pressure air
'Air pressure'

Levi (1976) informally classifies the semantic relations expressed in the CS into groups correlating to "degrees of productivity" (p.15), according to the frequency in which the relation types were found in various sources.[5] Possessive (2b), action/state (2c), purpose (2d), and attributive relations were found to be highly productive; a second group which includes content (2e), measure/container (2f) and classifying relations are somewhat less productive, while causal relations (2g) were found to be the least productive group. While it was found that there is variability in the use of the CS for expressing different relations, importantly for the purposes of this dissertation, possessive relations were found to be highly productive with the CS (though not semantically precise, it is common to refer to genitive nouns as possessives).

[5] Levi (1976) used as sources normative grammar books such as Avineri (1964) and Peretz (1968), random sequences of dictionary citations, translations of English data cited in Levi (1975) and lists drawn mainly from newspaper prose found in Reif (1968). (Note that these might not be fully representative of normal usage.)

However, most researchers on genitive constructions in Hebrew (e.g., Azar, 1976; Berman, 1978, 1988; Ravid et al., 1986) suggest that there are stylistic differences in the use of the CS and FS constructions, as mentioned briefly above: The CS is typically found in more formal expository style and in most written Hebrew, in marked contrast to speakers' preference for the FS in informal speech. Although this is the widespread belief, it is only partially confirmed by Ravid & Shlesinger (1995). They observe that the more formal the discourse, the higher the proportion of the CS (78% CS occurrences in written discourse compared with 58.3% CS occurrences in spoken discourse), and the lower the proportion of the FS (15% FS occurrences in written discourse compared with 39% FS occurrences in spoken discourse), but note that CS is more frequent than FS even in spoken uses in their data.[6]

Evidence from the acquisition of Hebrew complex nominals indicates that children acquire the CS construction relatively late: Berman (1986) shows that children are able to produce innovative CS nominals only at age 4-5, and only at the end of grammar school do children master most of the morphophonological alternations of the CS (this property of the CS will be discussed further in §2.3). In contrast, by age 3, children have already mastered the FS. This is expected given that the morphophonological properties of the FS are simpler than those of the CS: the FS is composed of two unbound nominals, as opposed to the CS, whose head is a bound form which undergoes morphophonological modifications on the free form.

Ravid & Shlesinger (1995) attempt to explain an apparent paradox: On the one hand, a common assumption made in the literature is that the CS is historically the source

[6] This raises the possibility that these corpus frequencies affect the processing of FS and CS differentially in listening and reading. Though it's worth bearing in mind, it is not easy to establish it in the present experiments.

from which the FS evolved, therefore the more basic form (e.g., Berman, 1978, 1988; Ravid et al., 1986). On the other hand, the proportion of use of the CS increases with the rise in level of formality, and the CS becomes productive relatively late in the acquisition process. They propose that there are two main semantic types of relations conveyed with genitive constructions in Hebrew: those that denote possession, and classifying nouns, where the second noun defines the first noun. They claim that the typical unmarked CS expresses a classifying relation: the CS is an endocentric complex nominal "with focus on the head which constitutes the semantic core of the [CS], while the modifying component delimits it in various ways" by reference to its content, substance, purpose etc. (p. 160; see, for example, (2d) above). Ravid & Shlesinger claim that the FS, on the other hand, is used as the default form to express the relation of possession (both alienable and inalienable). They argue that in possessive relations, there is a balanced share of semantic weight between the possessed and the possessing elements, and the FS expresses this more transparently by specifying each phonologically independent element separately and giving explicit expression to *šel* 'of', the possessive marker.[7] Thus, although the CS *can* be used to express possessive relations, they suggest that it is regarded as marked in that role. Ravid & Shlesinger (1995) support these claims with evidence that the major function of the CS in child language is to express classifying relations, rather than possessive relations.

In summary, as evident from the brief discussion above, a range of semantic relations may be expressed using the FS and CS constructions, and both constructions alternate fairly freely when used to convey the relation of possession, which is the focus

[7] According to Ravid & Shlesinger (1995), Levi (1976) and Berman (1978), among others, in Biblical Hebrew the CS form was favored for expressing possession. However, in Modern Hebrew, the CS is no longer considered the preferred option.

of this dissertation. Importantly, the expression of possessive relations is grammatical and feasible in both constructions; they are both familiar to adult native Hebrew speakers and are prevalent in ordinary use of the language, both written and oral. While the two forms may have different attachment preferences, the intermingling of CS and FS in the experiments designed for this dissertation is not unnatural and is not expected to have untoward consequences for the experimental outcomes.

2.2 Syntactic properties of FS and CS: Similarities and differences

It is relevant to consider the different syntactic properties of the FS and the CS to see whether they explain, for example, why these constructions are subject to different restrictions on definiteness interpretation (see below) and yet exhibit the same word order and the same compositional meaning. This section first presents the ways in which these two constructions are syntactically similar. It then discusses the differences between the FS and the CS by focusing on the unique properties of the CS. These peculiarities have put the CS construction under the syntactic microscope in the literature (Berman, 1978; Borer, 1996, 1999; Ritter, 1986, 1988, 1991; Siloni, 1996, 1997, 2001, 2003; Hazout, 1991, 1995; Longobardi, 1996, 2001; Shlonsky, 2004; Dobrovie-Sorin, 2000; among others), and it is possible to draw from this literature a number of distinguishing properties in order to characterize the two constructions and move toward a syntactic derivation for each.

FS and CS nominals share the same word order for their two constituents: both are strictly right branching, and the head (N1) precedes its associate (N2). Process nominals, which take the form of either an FS or a CS (e.g., *ha-harisa šel ha-cava* (FS) or

harisat ha-cava (CS, both: 'the destruction of the army'), exhibit a strict word order in which the head noun is followed by a logical subject and then a logical object, creating an N-S-O word order (Borer, 1996).[8] In fact, as noted in Siloni (1996), arguments of the process noun always appear post-nominally. In addition, in both FS and CS the head of the construction (the first noun) may take as its genitive phrase another complex nominal, which is itself composed of a head noun and a post-nominal genitive phrase. See (3a) for the FS and (3b) for the CS, which also demonstrate the similarity in meaning between the two complex nominals. Notice, also, that it is possible to mix the FS and CS when more than two nominals are involved, as in (3c,d).

(3) a. ha-bayit šel ha-axot šel ha-yeled
the-house of the-sister of the-boy
'The house of the sister of the boy'

b. beyt axot ha-yeled
house sister the-boy
'The house of the sister of the boy'

c. ha-bayit šel axot ha-yeled
the-house of sister the-boy
'The house of the sister of the boy'

d. beyt ha-axot šel ha-yeled
house the-sister of the-boy
'The house of the sister of the boy'

It is clear from the examples in (3) that the two constructions differ in at least three important respects. First, definiteness in the CS is marked only on the most embedded

[8] On the basis of the example in (3d) and the general relation between thematic relations and structural positions in CS process nominals, Siloni (1996, 1997) notes a thematic restriction on the ordering of arguments within the CS. She argues that N1 and N2 in non-process nominals exhibit some freedom in terms of their placement vis-à-vis each other. On the other hand, the base structure of process nominals is stricter, with the logical subject (N2) necessarily preceding the logical object (N1) in surface structure. This will be discussed further in §2.5 below.

noun, while in the FS it is marked on each noun separately. Second, in the FS, but not in the CS, the preposition *šel* surfaces between the two nouns. Third, the morphological realization of the word 'house' is different in the FS (*bayit*) and in the CS (*beyt*). The following paragraphs present in detail the differences between these constructions.

One of the most interesting distinctions between the FS and the CS is definiteness marking and its interpretation. While in the FS the definite marker appears on each noun separately, in the CS definite marking appears solely on the most embedded noun, and modification of the head noun (or other nouns in the sequence) by a definite article is not possible.[9] This distinction has been demonstrated in the examples in (3). The absence of a definite marker on all but the (linearly) last noun in the sequence gives rise to a unique property of the CS, which is commonly referred to as 'definiteness spread': the definiteness value of the CS as a whole is determined by the definiteness value of the most embedded noun.[10] Consider (4a), in which N2 is definite and the interpretation of N1 is rendered definite as well, as opposed to (4b) where an indefinite N2 renders the whole complex nominal indefinite. Crucially, an interpretation by which N2 is definite and N1 is indefinite is not available, and nor is a reading according to which N2 is indefinite but N1 is interpreted as definite.[11] In the FS, on the other hand, the definiteness

---

[9] In Hebrew, there is no indefinite article, thus the claims about definiteness marking refer to [+definite] marking only. However, it is possible to assume that the same is true for covert indefinite marking by making a comparison of the CS in Hebrew and in standard Arabic (but see footnote 11). As shown in Longobardi (1996), Standard Arabic manifests the ungrammaticality of marking the head N1 of the CS with an indefinite article: even when N2 is indefinite (marked by a case ending (tanwin)), it is not possible to mark N1 for indefiniteness by using a case ending on N1. Thus Standard Arabic presents a case in which indefiniteness spread is arguably attested, and provides support for extending indefiniteness spread to Hebrew as well.

[10] Dobrovie-Sorin (2000) argues that definiteness is a mono-valent feature, and therefore there is no (in)definiteness spread, but only definiteness spread (cf. footnote 9).

[11] But see Siloni (2003) for examples of discrepancies in definiteness values between N1 and N2 formed by non-prefixal determiners on N1, and discussion of these facts in the analysis of the two constructions below. See also Danon (2008) who proposes that while the two nouns of the CS share the same syntactic definiteness value, they are not necessarily interpreted as both definite or both indefinite. Here I limit the

value is strictly associated with the noun marked as such, and therefore a four-way definiteness marking is possible in the FS, as illustrated in (5), as opposed to 2-way definiteness marking in the CS. (Note that Hebrew does not have an indefinite article; the marking of indefiniteness is touched on in footnotes 9 and 10).[12]

(4) a. beyt ha-yeled
house the-boy
'The house of the boy'
* 'A house of the boy'

b. beyt yeled
house boy
'A house of a boy'
* 'The house of a boy'

(5) a. ha-bayit šel ha-yeled
the-house of the-boy
'The house of the boy'

b. ha-bayit šel yeled
the-house of boy
'The house of a boy'

c. bayit šel ha-yeled
house of the-boy
'A house of the boy'

d. bayit šel yeled
house of boy
'A house of a boy'

---

discussion to cases in which the definiteness value of N2 of the CS does determine the definiteness value of the entire comple nominal, as is typically the case.

[12] Berman (1978) notes that although all 4 variations in (5) are deemed grammatical by Hebrew grammarians, the 'mixed' forms (5b) and (5c) are far less common in actual usage, and are dismissed as ungrammatical by some speakers. These are the versions which are not available in CS. (5b) has a generic interpretation: Berman uses the example, *ha-para šel 'ikar* 'the cow of a farmer' with the meaning 'the cow of a farmer, rather than the cow of a shoemaker', or 'I need a farmer's cow, not a plastic cow'. In using (5c) the idea is more like "one of…" or a double type of genitive, similar to the English 'a cow of the farmer's' in the sense of 'one of the farmer's cows'.

Evidence that the CS is syntactically interpreted for definiteness in accordance with the definiteness value of N2 draws on three types of arguments. First, adjectives in Hebrew generally agree with the nouns they modify for definiteness as well as for gender and number. Thus when N1 is modified by an adjective, we expect the adjective to agree with it in definiteness. This prediction is borne out: in (6a), an overtly definite adjective modifies a covertly definite N1, while in (6b) the adjective agrees with N1 and both are indefinite. Second, Hebrew has an accusative Case marker, which only surfaces preceding definite nominals (7). Thus, we would expect the accusative Case marker *'et* to be realized preceding N1 when the CS is definite (8a) and not to surface when the CS is indefinite; compare (8b) and (8c). As is evident from a comparison of (7) and (8), a definite article on N2 renders the whole CS definite for the purpose of Case marking.

(6) a. beyt ha-yelda ha-gadol
house the-girl the-big$_{(masc)}$
'The big house of the girl'

b. beyt yalda gadol
house girl big$_{(masc)}$
'A big house of a girl'

(7) a. Roni ra'a 'et ha-bayit
Roni saw ACC the-house
'Roni saw the house'

b. Roni ra'a bayit
Roni saw house
'Roni saw a house'

(8) a. Roni ra'a *('et) beyt ha-yalda
Roni saw ACC house the-girl
'Roni saw the house of the girl'

b. Roni ra'a beyt yalda
Roni saw house girl
'Roni saw a house of a girl'

c. * Roni ra'a 'et beyt yalda
Roni saw ACC house girl
'Roni saw a house of a girl'

The third type of evidence for definiteness spread is noted in Siloni (2003). Siloni demonstrates that a definite CS appears in environments that only allow definite nouns, such as in subject positions in predicational sentences (9) or as topicalized elements (10), while in certain other environments the definite CS is excluded. For example, in existential contexts induced by the existential particle *yeš* (11), only indefinite CS nominals are possible.[13] The examples below are taken from Siloni (2003).

(9) a. tmunat ha-praxim (hi) xadaša
picture the-flowers (is) new
'The picture of the flower is new'

b. *tmunat praxim (hi) xadaša
picture flowers (is) new
'A picture of flowers is new'

(10) a. tmunat ha-praxim ra'iti ota
picture the-flowers saw(I) it
'I saw the picture of the flowers'

b. *tmunat praxim ra'iti ota
picture flowers saw(I) it
'I saw a picture of flowers'

(11) a. * yeš tmunat ha-praxim ba-xeder
there-is picture the-flowers in-the-room
'There is the picture of the flowers in the room'

b. yeš tmunat praxim ba-xeder
there-is picture flowers in-the-room
'There is a picture of flowers in the room'

[13] In the sentences in (11), the PP 'in the room' is potentially ambiguous, although intuition suggests that there is a bias towards a non-local attachment in this example, where 'in the room' describes the location of 'the picture'.

I now turn to another set of properties which are characteristic of the CS and differ from those of the FS. These properties relate to adjectival modification of constituents in the CS (Berman, 1978, among others). In contrast with the FS, the CS head N1 cannot be directly followed by an adjacent adjective modifying N1. An adjective modifying the head noun (N1) must follow the whole complex nominal, as shown in (12a). These facts are captured by the 'adjacency constraint' (Dobrovie-Sorin, 2002), which prohibits an intervention of any element between N1 and the maximal projection of N2. In the FS, on the other hand, adjectival modification of the head follows the head directly (12b). In addition, when both nouns of the CS are modified by adjectives, this results in a nested configuration (13a); no such effect is evident in the FS, where each noun may be followed by an adjective modifying it, similarly to adjectival modification of simple nouns. An interesting observation can be made regarding the insertion of quantifiers modifying N2, in what appear to be violations of the adjacency constraint (14). In fact, only pronominal quantifiers and numerals can appear internal to the CS. This is natural, however, under the assumption that N2 is a maximal projection higher than NP (presumably a DP). The constraint is therefore rather on the insertion of any material relating to N1 between N1 and the extended projection of N2.[14]

(12) a. beyt *(ha-gadol) ha-yalda ha-gadol
house$_{masc}$ (the-big$_{masc}$) the-girl$_{fem}$ the-big$_{masc}$
'The big house of the girl'

b. ha-bayit ha-gadol šel ha-yalda *(ha-gadol)
the-house$_{masc}$ the-big$_{masc}$ of the-girl$_{fem}$ *(the-big$_{masc}$)
'The big house of the girl'

[14] Notice also that N2 in (14) has a prenominal quantifier, not a definite article *ha-*, yet the whole CS is rendered definite. As pointed out by Siloni (2003), this suggests that *ha-* is not the only means by which a CS may be rendered definite (through definiteness spread).

(13) a. beyt ha-yalda *(ha-gadol) ha-yafa ha-gadol
house$_{masc}$ the-girl$_{fem}$ (the-big$_{masc}$) the-pretty$_{fem}$ the-big$_{masc}$
'The big house of the pretty girl'

b. ha-bayit ha-gadol šel ha-yalda ha-yafa
the-house$_{masc}$ the-big$_{masc}$ of the-girl$_{fem}$ the-pretty$_{fem}$
'The big house of the pretty girl'

(14) a. beyt kol yeled
house every the-child
'The house of every child'

b. beyt šloša yeladim
house three children
'The house of three children'

To sum up thus far, I have presented the major syntactic properties distinguishing the FS and the CS, arguing that the CS has unique characteristics which restrict the relation between its head and other elements. It seems plausible to assume that the unique characteristics of the CS are the result of the nature of its head (N1), which is restricted in several ways. This is evident from the fact that the head cannot be directly modified by a definite article or an adjacent adjective, and must participate instead in definiteness spread and resort to a nested adjectival configuration. I now turn to another unique property of the CS, the status of its head as a phonologically reduced (construct) form.

2.3 The phonological properties of the CS

It is widely assumed in the Hebrew literature that the CS is a single prosodic word (see, for example, Borer, 1988, 1996, 1999; Siloni, 2001, 2003; among others). The phonological properties of the CS manifest the status of the CS as a single prosodic word which follows the stress assignment patterns of simple nouns in Hebrew, as will be

argued below (see also Chapter 3, §3.3 for further discussion of lexical and rule-based stress in the Hebrew nominal system).

An early discussion of the phonology of the CS is Gesenius (1910)' characterization of the CS in Old Hebrew. The CS nominal was the only construction available in Biblical Hebrew to convey genitive relations (Levi, 1976; Berman, 1978; Ravid & Shlesinger, 1995). This early, seminal work on the grammar of Biblical Hebrew describes the CS in Biblical Hebrew, but it seems that the unique properties of the CS in Modern Hebrew were carried over from its form in Biblical Hebrew.[15] As can be inferred from Gesenius (1910), the definiteness facts as well as the modification restriction, as described above, are already attested in the CS in Old Hebrew. Phonologically, Gesenius (1910) characterizes the CS as "a close connexion (or interdependence) of the *Nomem regens* [the head in the CS] and the *nomen rectum* [the genitive phrase]." (p. 247). While the morphophonological form of the modifying noun remains entirely unchanged, it is claimed that the close combination of the two nouns results in placing the main stress on the second noun (N2); as a result of the weaker tone on N1, N1 also undergoes further phonological changes. Gesenius notes that the changes affect vowels more often than consonants: vowels in N1, which a phonological rule requires to be lengthened when positioned in or prior to a stressed syllable, necessarily undergo vowel reduction and

---

[15] Making analogies between Modern Hebrew and Old Hebrew should be done with caution, however, especially when stress and tone are involved. Although the two varieties share many properties, important systematic differences arise. For example, the stress on verbs in Old Hebrew is placed on the final syllable (i), with reduction of the initial vowel, while the stress in Modern Hebrew is on the penultimate syllable in the example in (ii), and no vowel reduction occurs (see further discussion of stress patterns in Chapter 3 §3.3).

(i) gmartém
(ii) gamártem
Both: finish (2, pl, past tense)

However, the phonological and syntactic behavior of the CS in both varieties seems highly similar, and it is helpful to draw from this literature on the nature of the CS, especially because the literature on the phonology of the CS in Modern Hebrew is so limited.

become shorter in the CS. The reduced form of the head of the CS has resulted in its name 'dependent' in Hebrew.

In contemporary discussions of the CS, the head of the CS (N1) is commonly described as a bound morpheme. Berman (1978) discusses the fact that any noun in Hebrew has potentially two surface forms: free and bound. The phonemic representation of the two forms is identical, according to Rosen (1977), but there are subphonemic differences which are the result of affixal alternation. Thus, N1 of the CS appears in its bound form due to the affixation of N2: it is stripped of "its proper stress and all its vowel phonemes are materialized as stress-remote" (Rosen, 1977:141). This accounts for the appearance of processes such as *schwa* (ə) replacement (e.g., *qacir* 'harvest' becomes *qcir* in *qcir xitim* 'wheat harvest', see also the main stress in *bayit* in (15a), and the reduced form *beyt* in (15b) and (15c)).[16] As a bound form, N1 cannot stand on its own as an independent phonological word and must be followed by DP2. The lack of stress on N1 in the CS appears to give rise to the same morphophonological modification which occurs preceding pronominal clitics, where the head noun appears in the bound form with a pronominal element suffixed (stress is on the last syllable, which contains the pronoun clitic). The example in (15a) presents the Hebrew noun for 'house' in its free form; (15b) shows the same noun in its bound form within the CS; (15c) shows the bound form of this noun with a pronominal clitic suffixed to it (the word *bayit* 'house' in its bound and free form appears in bold).

[16] Rosen (1977) comments on the similarity in phonological alternation between the bound form of the CS and other nouns which undergo affixation, using the example of plural alternation. He notes that the same type of alternation can be found in the relation of singular to plural forms of masculine nouns, where the additional suffix of the plural form attracts the accent and creates the conditions for the alternation discussed, e.g., *davar* 'thing' vs. *dvarim* 'things'.

(15) a. **báyit**
house
'A house'

b. **beyt** ha-yéled
house the-boy
'The house of the boy'

c. **beytí**
house-my
'My house'

Berman (1978) outlines two main types of processes, vowel reduction and final /t/ deletion, which are the result of stress reduction on N1 of the CS.[17] She notes that these alternations occur in any Hebrew noun to which a suffix is added (e.g., the plural or possessive) and that N2 of the CS has just the same effect, phonologically speaking, as any other affix. Vowel reduction is generally characterized as the deletion of the penultimate non-high vowel of the noun in the presence of a stressed suffix; an identical set of rules applies to the head noun of CS. The examples in (16) (taken in part from Berman, 1978) demonstrate this: (16a) presents the noun in its free form; (16b) shows the same noun in its bound form when plural marking is suffixed to it; (16c) is an example of the bound form of the noun when it is N1 of a CS. The stress shift and deletion do not occur when the noun in question is modified by an adjective (16d), because it appears in its free form.

[17] The majority of Hebrew nouns have the same surface form in the free and the bound forms (e.g., nouns with the feminine plural ending *-ot* take the same form whether free or bound). A number of nouns have different bound and free forms for historical reasons which are only partially determined by regular phonological processes still productive in Modern Hebrew (e.g., nouns such as *xaver* 'friend' with the masculine plural ending *im* (*xaverim* 'friends') takes the special ending *ey* in the construct: *xavrey yaldut* 'childhood friends'). Several alternations form part of the more or less idiosyncratic knowledge the native speaker has of certain words in the language. E.g. *bayit* 'house' has a non-diphthongized bound stem *beyt*, which occurs both before suffixal endings and as N1 of the CS.

(16) a. **nicaxon**
victory (free form)

b. **nicxon**ot
victory (bound form) + plural suffix *ot*
'Victories'

c. **nicxon** ha-cava
victory the-army
'The victory of the army'

d. **nicaxon** moxec
victory crushing
'A crushing victory'.

A second type of alteration is final /t/ deletion. Hebrew has a general rule deleting word-final /t/ when preceded by the vowel /a/, under certain quite complex constraints (see Berman, 1978:255 for a detailed description). Thus the underlying representation of the Hebrew word for 'bride' in (17a) has a final /t/ which is deleted due to the /t/ deletion rule (see, for comparison, the Hebrew word for 'store', which does not undergo /t/ deletion as is evident from the surface realization of the final /t/ in the free form (17a)). When an enclitic suffix is added to the word, /t/ is no longer in final position and therefore the underlying /t/ is vocalized (17b). Berman (1978) claims that N2 of the CS has the same effect as an enclitic suffix: The fact that /t/ is retained in the CS examples indicates that there is no word boundary between the head N1 and the adjunct N2 of the CS, suggesting that they constitute a phonological word.

(17)	a.	**kala** bride	cf.	**xanut** store
	b.	**kalat**i bride.mine 'My bride	cf.	**xanut**i store.mine 'My store'

c.	**kalat** ha-xatan bride the-groom 'The bride of the groom'	cf.	**xanut** xašmal store electronics 'A store of electronics'

Prince (1975) argues along similar lines that the phonological properties of the CS in Tiberian Hebrew (i.e., the reduced stress on N1 which is subordinate to the main stress on N2 and the phonological modifications on N1) can be seen as following directly from the stress pattern of all nouns, under the crucial assumption that the entire CS forms a single phonological word. Prince (1975) argues that the string of nouns in the CS is analyzed by the phonology as a single word, and therefore the rules of the phonology apply to the CS as a whole, not to the individual nouns within it. The phonological processes discussed above occur as consequence of the absence of a boundary intervening between the nouns within the CS, due to its status as a single phonological word.

In her study of the stress system of Modern Hebrew nouns, Bat-El (1993) includes compound nouns; a full discussion of compounds is presented in §2.4 below, but since Hebrew compounds are phonologically similar to CS nominals, the phonological aspect of this construction is discussed here. Bat-El notes that "the main stress of the rightmost element is the main stress of the compounds, and secondary stress falls on the main stress of the first element." (Bat-El, 1993:206). For example, the two free forms *xatúl* 'cat' and *rexóv* 'street' each carries a main stress on the final syllable of the noun; when the two forms are used in a compound/CS, the final syllable of 'cat' carries a secondary stress: *xatùl rexóv* 'street cat'. Bat-El (1993) therefore proposes that in addition to stress assignment rules which assign main stress to each member of the compound/CS separately, another stress rule assigns stress to the head of the compound/CS, which is

assumed to be phonologically "an unbounded right-headed constituent" (Bat-El, 1993:206). The compound/CS stress assignment rule applies after other stress rules in the metrical grid she proposes for the Hebrew nominal system. She illustrates the application of the rules for the compound *medinàt Israél* 'Israel state' ('State of Israel'), see example (18) (Bat-El's (1993):207, example (30)), and suggests that phonologically, the head of the compound/CS is N2 (See Chapter 8 §8.3 for implications of this assumption).

(18)

	*	Level 3 (compound)
(*	*)	Level 2 (word)
(*)	(*)	Leve1 (foot)
(* * *)	(* * *)	Level 0 (syllable)
me di nat	yis ra el	

Borer (1988) offers a test in support of the word status of the CS[18] as opposed to the FS by the evident ungrammaticality of the CS in the case of extraction of N2 out of the CS (19b) (example from Borer (1988):52; [e] corresponds to the gap created by the WH operation).

(19) a. lo barur šel mi_i ran xaver $[e]_i$ FS
not clear of who Ran friend [e]
'It is not clear whose friend Ran is'

b. * lo barur mi_i ran xaver $[e]_i$ CS
not clear who Ran friend [e]
'It is not clear who Ran is a friend of'

[18] Borer argues that this is evidence for the fact that the CS is a *syntactic* word, giving rise to extraction restrictions. (But see a different interpretation of these facts in Siloni, 2003.)

In sum: the CS is widely argued in the literature to constitute a prosodic word. This conclusion is based on the unique phonological and syntactic properties it manifests, namely, stress reduction on the head N1 as well as phonological modification of this head, which have been shown to apply to other nouns in Hebrew when they undergo affixation (see further discussion of Hebrew stress assignment in Chapter 3, §3.3). The experimental data in this dissertation can be used to provide further support for the status of the CS as a single phonological entity.

2.4 Hebrew compounds

A full discussion of complex nominals in Hebrew cannot be complete without mentioning Hebrew (lexical) compounds. The Hebrew noun compound is similar in many ways to the CS nominal and its characterization is subject to an ongoing debate in the literature (see below). Compounds seem on the surface to share with the CS some of its special properties; see examples of the CS (20a) and the compound (20b) below. Several similarities are immediately observable: First, the head of the construction appears in the bound form and undergoes the same morphological alternations resulting from stress reduction: in both examples in (20), N1 is in its bound form *beyt* (cf. the free form *bayit* 'house').[19] Second, in both CS and compound cases, only the final, most embedded noun in the sequence is marked for definiteness; the definite article cannot be affixed to the head of the construction directly (though an outer modification of the

[19] Along similar lines to the explanation proposed for the CS in §2.3 above, Bolozky (1982) claims that the stress pattern in compounds, specifically that of N1, should be attributed to the distribution of subsidiary (secondary) stress. He argues that the constituent words of which each compound (e.g., *'orex din* 'editor law', meaning 'lawyer') is composed have merged into a single lexical unit, thus explaining the reduced stress on N1 of the construction.

compound as a whole by a preceding definite marker is possible for most or all compounds in colloquial Hebrew, as will be discussed below)[20].

(20) a. beyt ha-yeled
house the-boy
'The house of the boy'

b. beyt ha-sefer
house the-book
'School'

In addition, both the CS and compounds may be iterative in the sense that they can serve, as a whole, as the complement (N2) of a construct noun (N1), thus forming a chain of nouns in which all but the last noun are in the bound form (21). Also, they may not be modified internally by an adjective, i.e., no adjective may intervene between the two nouns in (22).

(21) a. tmunat beyt ha-yeled
picture house the-boy
'The picture of the house of the boy'

b. tmunat beyt ha-sefer
picture house the-book
'The picture of the house of the book'
(lexicalized meaning: 'The picture of the school')

(22) a. * beyt ha-gadol ha-yeled
house the-big the-boy
'The big house of the boy'

b. * beyt ha-gadol ha-sefer
house the-big the-book
'The big house of the book'
(lexicalized meaning intended: 'The big school')

[20] In the language of more conservative speakers, there are some compounds that might resist it, due to the prescriptive grammar of the Academy of the Hebrew Language. However, in current speech of younger people, it is probably true that all compounds allow an outer article.

Although they share some properties, the CS and compounds also have a number of distinguishing properties. All of these differences are useful to distinguish compounds from the CS examples used in the materials for this dissertation. The FS rephrasability criteria (see below) was in fact used as a pretest, but intuitively every one of these criteria was satisfied by the CS sentences used for this dissertation. The most obvious distinction between CS and compounds is in terms of their semantic compositionality: while the meaning of the CS is compositional and is derived from the meaning of the two nouns, the meaning of compounds is often (though not always, and to different degrees) semantically opaque; thus many compounds are considered lexicalized or frozen expressions. E.g., the CS in (20a) is composed of the meaning of N1 *beyt* ('house') and N2 *yeled* ('boy'); in contrast, the compound in (20b) carries the meaning 'school' which is unpredictable and cannot be straightforwardly derived from the meaning of the individual semantic components *beyt* ('house') and *sefer* ('book').[21] In fact, for many lexicalized compounds, the literal meaning (e.g., 'the house of the book') is implausible and even nonsensical. Rosen (1977) notes that this semantic property may be used to decide whether compounding has taken place or not. Note, though, that noncompositionality is a sufficient but not a necessary condition for compound status. Another semantic property distinguishing compounds and CS nominals is the severe restriction in compounds (but not in CSs) on the replaceability of any one of the components by members of its substitution class: in (20a), *binyan* 'building' may replace

[21] Peretz (1968) argues that examples such as *beyt sefer lerxiva* 'horse back riding school' and *beit sefer letfira* 'sewing school' demonstrate the fact that the meaning of the individual nouns composing the compound ('house' and 'book') have weakened or disappeared and a new term was created. He claims that although in the original combination 'book' served a semantic function (i.e., describing the type of learning characterizing the place, i.e., learning books), the examples above do not involve learning related to 'book'; rather, the words 'horse back riding' and 'sewing' are argued to replace the function of 'book' in the original combination.

beyt ‘house’, or *‘iš* ‘man’ may replace *yeled* ‘boy’; on the other hand, in *beyt sefer* ‘school’ neither noun is replaceable.

Compounds differ from CS nominals syntactically in several important respects, which have been highlighted in several works in the literature (e.g., Baker, 1998; Berman, 1978; Berman & Ravid, 1986; Bisetto & Scalise, 1999; Borer, 1988; Sadock, 1998; among others). Most of these distinctions are the result of restrictions on N2 in compounds. The first distinction concerns coordination of N2 (under head deletion). While it is possible to coordinate N2 of a CS (23a), this is not possible with compounds (23b). When N2 of a compound is coordinated, the lexicalized meaning is lost (compare with the lexical meaning of each compound when not coordinated, in (23c). The compound examples in (23) through (29) are all grammatical; what is demonstrated here is the loss of the compound (lexicalized) meaning. The resulting interpretation is compositional, quite similar to that in the equivalent CS examples, but it is often implausible in the case of compounds, once the lexicalized meaning is lost.

(23) a. beyt ha-yeled ve-ha-yalda
house the-boy and-the-girl
‘The house of the boy and the girl’

b. beyt ha-sefer ve-ha-xolim
house the-book and the-patients
‘the house of the book and the patients’
* ‘The school and the hospital’

c. beyt ha-sefer beyt ha-xolim
house the-book house the-patients
‘School’ ‘Hospital’

Another well observed distinction between the CS and compounds is pronominal reference. While it is possible to refer to the head of the CS by using a pronoun, as in (24a), such pronominal reference is impossible in the compound example (24b). Sadock (1998), however, rules this out as a general test for compoundhood; he claims that pronominal reference in English is acceptable for synthetic compounds (e.g., *Gingrich supporters think he is brilliant*) and for anaphoric compounding (e.g., *Bill tried to seek self-fulfillment*). Though his point may be valid for English, Borer (1988) claims that this is not the case for the Hebrew compounds, for which pronominal reference is ungrammatical under the lexicalized meaning; see examples in (24), from Borer (1988).[22] Note also that N2 of compounds is restricted in that its number specification cannot be changed without the compound losing its lexicalized meaning (and it is not related to the number specification of the compound as a whole; plurality in compounds is marked exclusively on N1, e.g., *batey sefer* 'houses book' meaning 'schools')[23]. The CS in contrast is not restricted in that way: number marking can be assigned to each of the nouns separately.

---

[22] There seems to be a sense of variability in acceptability of pronominal reference in compounds that are not highly lexicalized such as *xadar šena* ('room sleep', lexicalized meaning 'bedroom') and *xadar 'orxim* ('room guests' lexicalized meaning 'guest room'). Consider the semantically acceptable example in (i) below:

i. be-bayit ze yeš šney xadrey šena ve-'exad le-'orxim
in-house this there-are two rooms sleep and-one for-guests
'This house has two bedrooms and one guestroom'.

As with the head (N1) of example (20b) in the main text (*beyt* 'house'), N1 in this example (*xeder* 'room') is often used as the head of compounds. This might suggest that the relation of purpose (used in the example here to express the purpose of different 'rooms') is more likely to become lexicalized than other relations expressed by the CS.

[23] In colloquial Hebrew, some highly lexicalized compounds are pluralized by adding the plural suffix to the compound as a whole (*orexdinim*), instead of the prescribed requirement that the plural suffix be marked on the first noun of the compound (*orxey din*); see further discussion of this property below.

(24) a. hu bana li šney batey 'ec ve-'exad mi-plastik
he built for-me two houses wood and-one from-plastic
'He built me two wood houses and one from plastic.'

b. ?? hu bana lanu šney batey xolim ve-'exad le-zkenim
he built for-us two houses sick(pl) and-one for-old(pl)
(cf. beyt xolim; beyt zkenim
house sick(pl) 'hospital' house old(pl) 'retirement home')

CS and compounds also differ in the adjectival modification of N2. While it is possible to modify N2 by an adjective in the CS, this is not possible for compounds. Since adjectival modification in both CS and compounds always follows the N2, regardless of whether the adjective modifies N1 or N2 (see above), the construction is ambiguous in cases where the two nominals share identical features. However, when N1 and N2 do not share the same gender and number features, the adjective which follows N2 agrees in gender and number features with only one of the two nominals, and a distinction between CS and compounds arises. In compounds, adjectival modification of N2 alone is not possible while maintaining the lexical meaning (25b), while for CS nominals it is possible to modify N2 to the exclusion of N1 (25a).

(25) a. beyt ha-yalda ha-xamuda
house the-girl the-sweet
'The house of the sweet girl'

b. beyt ha-xolim ha-zkenim
house the-sick the-old
'The house of the sick old people' (*beyt ha-xolim* means 'hospital')

N2 in a compound is also restricted in its ability to be modified by cardinal numerals, which in Hebrew take the construct form when preceding the noun they modify. Once a numerical is inserted, the compound loses its lexicalized meaning (26b),

demonstrating that N2 is not accessible to any modification when part of a compound. This restriction does not apply to N2 of the CS (26a).

(26) a. beyt šney ha-yeladim
house two the-children
'The house of the two children'

b. beyt šney ha-sfarim
house two the-books
'The house of the two books'

Another distinction between the CS and compounds does not involve manipulations on N2; rather it is an internal/external argument distinction, which is manifested specifically in process nominals. As Borer (1988) points out, while a logical object can assume N2 position in both the CS and compounds (27a,b), a logical subject can assume N2 position only in the CS (28a) but not in compounds (28b) (examples from Borer, 1988:50-51).[24]

(27) a. axilat ha-tapu'ax al-yedey dan
eating the-apple by Dan
'The eating of the apple by Dan'

b. giluy ha-'arayot al-yedey ha-'av
exposure the-genitals by the-father
'Incest committed by the father'

(28) a. axilat dan 'et ha-tapu'ax
eating Dan ACC the-apple
'Dan's eating of the apple'

[24] Baker (1998) comments extensively on the restrictions on the use of process nominals in compounds and argues that they follow from independent syntactic principles such as the theta criterion, transitivity of the verb and other principles. However, his account is based on evidence from different languages and does not explain the distinctions between Hebrew CSs and compounds in that respect. Notice, for example, that both examples in (28) are nominals derived from transitive verbs, and yet the first is grammatical while the second is clearly ungrammatical under a compound reading.

b. giluy ha-'av 'et ha-'arayot
exposure the-father ACC the-genitals
'The father's exposure of the genitals'
(* Compound meaning: 'the father's committing incest')

A clear difference exists between CS and compounds such that only CS nominals, but not compounds, may be paraphrased into a FS (Berman & Ravid, 1986). When compounds are substituted by a FS, their lexicalized meaning is lost. The examples in (20) are repeated here, each followed by an FS substitution.

(29) a. beyt ha-yeled
house the-boy

a'. ha-bayit šel ha-yeled
the-house of the-boy

both: 'The house of the boy'

b. beyt ha-sefer
house the-book
'school'

b'. ha-bayit šel ha-sefer
the-house of the book
'The house of the book'
(* compound meaning: 'the school')

The final distinction between the CS and compounds comes from data from Colloquial Hebrew. As briefly mentioned above, only in compounds can the definite article appear outside the construction (an outer article), in alternation with the traditional (formal) position of definite modification, internal to the construction.[25] The outer definite article that precedes the compound renders the whole complex nominal definite

[25] The possibility of outer instead of inner definiteness modification of this construction is briefly mentioned by Ritter (1988: footnote 6), although it is not incorporated into the account proposed there. It is also mentioned in Siloni (2003). Neither proposes a syntactic analysis for this relatively new development in the language.

(30a), as does an internal article affixed to N2 (30b). In the case of the CS, the outer article is prohibited (31a); the definite article can only appear attached to N2, as demonstrated in (20) above, repeated here as (31b). Note that when the definite article precedes the compound, there cannot be an additional definite marker on N2, and N2 cannot be interpreted as indefinite (30c).[26]

(30) a. ha-beyt sefer
the-house book
'The school'

b. beit ha-sefer
house the-book
'The school'

c. * ha-beyt ha-sefer
the-house the-book
'The school'

(31) a. * ha-beit yeled
the-house boy
'The house of the boy'

b. beit ha-yeled
house the-boy
'The house of the boy'

The use of an outer article in compounds suggests that the compound is analyzed as a single noun, modified by a prefixed article in the same way that definite modification is marked on simple nouns in Hebrew: In (30a), the article precedes the whole nominal, similar to the position of articles in simple nouns (e.g. *ha-me'amen* 'the coach'). In fact,

[26] Although the use of the prefixed article construction is restricted to colloquial registers, it seems to be becoming more frequent, especially when used after prepositions such as 'to', 'in' and 'from' (e.g., (i) below), where the preposition (*be-*) and the definite article (*ha-*) typically contract (as *ba-*) and cliticize into N1.

i. niš'arti ba-beyt sefer
I-remained in.the-house book
'I remained in the school'

most of compounds allow an initial definite article in colloquial Hebrew (additional examples include *ha-yom huledet* 'the birthday', *ha-'orex din* 'the-lawyer', *ha-mexonat kafe* 'the coffee maker', and numerous others), though CS nominals that have recently been lexicalized or are in the process of lexicalization (i.e., becoming compounds) are less acceptable with an outer modification.[27] The outer modification used for compounds in colloquial Hebrew might be viewed as the emergence of a new compounding form, a manifestation of lexicalization constituting a relatively recent development in Hebrew grammar, which differs in its grammatical behavior both from productive CS and compound forms.

The differences between CS and compounds have led some researchers to identify them as two distinct constructions with different derivations. Other analyses, in contrast, maintain that these constructions are fundamentally the same. In what follows, I briefly review the types of proposals made in the literature for the derivation of compounds and how it is related to the derivation of CS nominals. For space reasons I will not present a detailed evaluation of these proposals, which are not the focus of the experimental study reported in the chapters that follow.

Borer (1988) argues for a unified Word Formation component of the grammar, and suggests that the CS and compounds are two instances of phonological words that are derived through word formation operations that are similar, but take place at different

[27] An additional development in colloquial Hebrew as well as in colloquial writing (e.g., email) is that many compounds are contracted to form a unified phonological form. Thus, the last consonant of N1 and the first consonant of N2 are contracted into a single consonant. For example, *beyt sefer* becomes *beycefer* ('school'); *yom huledet* becomes *yomuledet* ('birthday'), and is presented orthographically as a single word. A parallel process seems to have occurred in nouns such as *kaduregel* 'football' (from *kadur regel* 'ball foot') and *migdalor*'lighthouse' (from *migdal 'or* 'tower light'), which, according to Rosen (1956) lack a recognizable N2 in the orthography, but are structurally composed of two nominals. In addition, as briefly mentioned above, some compounds in colloquial Hebrew can have plural endings marked on the entire compound (e.g., *'orexdinim* vs. *'orxey din* 'lawyers') and not on N1 (though this form is relatively new and uncommon).

levels of representation (one in the syntax, the other in the lexicon); this is also assumed in Ritter (1986, 1988). It is related to the view taken by Berman (1978) and Dobrovin-Sorin (2000), who do not treat compounds and CSs as different constructions but rather assume that compounds constitute simply a frozen version of the productive CS. In a similar vein, Bisetto & Scalise (1999) adopt the tests for compoundhood suggested in Ten Hacken (1999) and claim that the CS is formed in the syntax, while compounds are formed in morphology. In contrast to the above accounts, which assume two different derivations for the two constructions, Sadock (1998) and Baker (1998) both hold the view that the CS and compounds should be analyzed as two manifestations of the same construction. However, here the similarity comes to an end: while Sadock claims that both the CS and compounds are formed in morphology, Baker takes the view that the formation of both is in the syntax.

Berman (1978) notes that the precise characterization of compounds "is not a simple matter, as it seems to be a lexical question of conventional usage rather than of the specific semantic or syntactic relations obtaining between the two nominals in each case. We would suggest that these compounds need to be viewed essentially as 'idioms'" (Berman, 1978: 235). She claims that there is an equally wide range of relations that can be expressed by CS and compound forms, and therefore it is not possible to establish any general motivated reason or special properties that would lead to the creation of compounds. The idiomatic view of compounds taken by Berman is that, like idioms, compounds start out diachronically as productive CS forms; at some point, they acquire a special meaning in some specialized contexts (by ad hoc inferences). After consistent use

of the complex nominal with that contextually adapted interpretation, the innovative interpretation gets lexicalized, i.e., enters the lexicon.[28]

In a questionnaire study of compounds and productive CS forms, Berman & Ravid (1986) claim that while both compounds and CS forms are created in the syntax, it is possible to place compounds and CS nominals on a scale according to the level of lexicalization they have undergone. On one end of the scale are compounds which are frozen referential units each constituting one single lexeme, and retrieved from the lexicon as autonomous, unanalyzed lexical units (similar to other nouns in Hebrew); on the other end of the scale are productive CS nominals which are analytic forms conceived as two lexical units and are analyzed by regular semantic/pragmatic processes. Berman & Ravid argue that the productive CS forms are constantly in a process of being drawn to the lexicon and slowly becoming more lexicalized.

The conditions under which a CS becomes 'fossilized' into a compound are not clear. Although, as mentioned above, establishing any general, motivated reason for the lexicalization of certain CS forms is difficult since an equally wide range of relations can be expressed by both fixed and productive CS, Berman (1978) and Berman & Ravid

---

[28] It should be noted that Hebrew compounds share many properties with idioms; in fact, most of the properties described for compounds can be explained as restrictions on idiomatic expressions. For space reasons, however, further investigation into an analysis of compounds as idiomatic expression is not pursued here. The general question of where idioms are formed is under debate. The traditional view of idioms assumes that idioms are stored units in the lexicon. But this view raises objections, in particular since it does not account for clausal idioms that exist only as syntactic transforms (e.g., *you have no leg to stand on*, 'you have no justification/chance of success'), which need to undergo syntactic operations and hence would appear to be post-lexical. An alternative view is proposed in the framework of Distributed Morphology (see Halle & Marantz, 1993; Marantz, 1997). In this framework, there is no component of the grammar specifically designed for word formation. Instead, syntax is responsible for both word and phrase formation. According to this view, idioms are assigned a structure by the syntax, but at LF the 'Encyclopedia' (which contains entries that relate vocabulary items to idiosyncratic meanings) associates a special non-compositional interpretation with the idiom, overriding the compositional meaning that arises if no special encyclopedic interpretation is associated with it. McGinnis(2002) gives an analysis of idioms in this framework. (On the other hand, see Horvath & Siloni, to appear, who claim that Hebrew verbal idioms cannot be stored in the Encyclopedia.) I leave the interesting question of the locus of idiom-formation open, since it does not have direct consequences for this study.

(1986) note that certain nouns seem particularly productive as the head (N1) of compounds. This is usually the case with nouns representing large semantic categories, e.g. *bayit* 'house' (*beyt sefer* 'school', *beit kneset* 'synagogue', *beit xolim* 'hospital', *beyt xarošet* 'factory' and so on), *xeder* 'room' (*xadar šena* 'bedroom', *xadar 'oxel* 'dining room' among many other), *mexona* 'machine' (*mexonal gilu'ax* 'razor, shaver', *mexonal ktiva* 'typewriter', *mexonat tfira* 'sewing machine' etc.). The fact that the head is frequent as N1 in this construction, it is argued, makes the combination more likely to be lexicalized. In addition, Berman & Ravid (1986) mention that some semantic relations seem to become more strongly lexicalized than others. The former include container relations (*zer praxim* 'bouquet') and process relations (*yošev roš* 'chairman') as well as purpose relations (*mexonat kvisa* 'laundry machine'), material (*ugat šokolad* 'chocolate cake') and place (*migraš tennis* 'tennis court'). Finally, the pragmatic context in which the CS is used (culturally and in specific circumstances) may foster lexicalization in some cases (e.g., *'aliat mexirim* 'price rise'*, niyarot erex* 'securities (financial)'*, necigat miflaga* 'party representative'), yielding a compound. The more limited the pragmatic context, it is claimed, the more fixed the expression.

One of the important findings in Berman & Ravid's (1986) study is that there was a statistically significant high level of agreement among participants in their judgments of each item.[29] The finding that some items were judged by every participant categorically

[29] Berman & Ravid's (1986) study presented 30 nominals (one third were antecedently judged as being highly lexicalized CS (compounds); another third were regular CSs and were not considered as lexicalized; the last third were cases on which there was no consensus between expert judges as to the level of lexicalization). The 108 participants were asked to evaluate each item according to five definitions, representing a descending scale of lexicalization (i.e., the expression sounds like (a) one word (b) a known and accepted combination (c) a combination that makes sense (d) a random combination never encountered before (e) an impossible combination in Hebrew. In addition to the high level of agreement between participants, the results showed that Hebrew speakers are tolerant for new and unfamiliar CS: even unlikely combinations were not rejected across the board. Finally, it was found that items with borderline judgment

as lexical items led Berman & Ravid (1986) to conclude that which items are lexical compounds is part of a native speaker's knowledge of Hebrew. Interestingly, agreement was most strongly manifested for those items considered by the experimenters to be compounds. This finding relates to Berman's (1978) view of compounds: "the constraint [under which an item becomes lexicalized] is purely a lexical one, in the sense that part of the idiosyncratic information a speaker has about the vocabulary of his language includes a listing of all and only those noun compounds which are essentially fixed or inseparable in Hebrew" (Berman, 1978:236).

To sum up, the discussion of CS and compound nominals has shown that although they share several unique properties, compounds are syntactically and semantically restricted in ways that productive CS nominals are not. As was found in Berman & Ravid's (1986) study and also intuitively, native Hebrew speakers seem to be in agreement on what constitutes a compound and what does not. As mentioned above, for purposes of this experimental research, the exact linguistic derivations of CS and compounds are less important than the existence of diagnostic properties by which they can be distinguished. The CS materials I constructed for the experiments reported in this dissertation were carefully chosen such that they are not restricted semantically or syntactically as compounds are. All of the CS used in the experiments reported here necessarily can be substituted by the FS (as is required by the nature of the investigation), are productive and semantically compositional, allow pronominal reference to N1, number marking can be assigned to each of N1 and N2 separately, cannot use an outer

---

between 'one word' and 'random combination' were those combinations created by free syntactic operations which represent semantic relations most typical of the CS (e.g., container and process relations); It is suggested that the presence of so many CS complex nominals which bear such semantic relations makes the emergence of other CS items with such relations plausible (Levi 1976).

modification of the definite article and there are no syntactic restrictions on N2 (N2 can be coordinated, can change the number specification, can be modified by an adjective or cardinal numerals, and can be realized as a logical subject in process nominals). These properties are not available for compounds, as discussed above. Therefore, it is safe to conclude that the CS nominals used as materials in this dissertation are not fixed or lexicalized, but rather are productive CS forms.

2.5 The syntax of Hebrew complex nominals

I now turn to the syntactic derivations of the FS and the CS. The FS has not received much attention in the literature, perhaps because its analysis has seemed quite straightforward. The FS is assumed by many researchers (for example, Borer, 1996, 1999 and Shlonsky, 2004) to be analogous to the possessive construction in English. Others treat the genitive phrase of the FS as an adjunct (Dobrovie-Sorin, 2000; Ritter, 1986, 1988). Yet other accounts offer an analysis of the FS with a base-generated structure identical to that of the CS in terms of the position in which arguments are inserted (Ritter, 1991; Siloni, 1996,1997), and derive the differences in their properties from the need of the N2 phrase to be assigned structural genitive Case in the CS, but not in the FS, as will be presented below.

The CS, on the other hand, has always been considered a challenge for analysis due to its unique properties and has therefore been the topic of detailed discussions, which continue to date. It is possible to divide the literature on the CS into four major groups of syntactic accounts: (1) the ‘incorporation’ account (Borer, 1996, 1999) argues that the CS exhibits word properties, not only from a phonological perspective, which is

theoretically undisputed, but also syntactically. It is assumed that the CS becomes a word at a stage that affects the interpretation,[30] and that it is the result of incorporation. (2) The head-movement account was originated in Ritter (1986, 1988, 1991) and has been adopted by many researchers (albeit with some minor differences in motivation) including among others Siloni (1996, 1997), Hazout (1991), Longobardi (1996, 2001) and references therein. It proposes that the unique properties of the CS are the result of the raising of N1 to D. This dissertation follows this approach, and it will be discussed in detail below. (3) The phrasal movement account (Shlonsky, 2004) argues for a generally similar derivation for both simple nouns and CS nominals. Based on the order of NP modifiers in different Semitic languages (which are claimed to have similar NP structure), and assuming that modifiers are base-generated to the left of NP and stay in situ, the different word order realizations are assumed to be the result of phrasal movement of the NP across its modifiers.[31] Finally, (4) the account by Dobrovie-Sorin (2000, 2002) does not involve any movement, and instead derives the surface order of the CS on the assumption that the extended projection of N2 is base generated in specifier position of D, which is right-peripheral.[32]

Non-syntactic accounts also exist in the literature on the CS: There is a proposal for the formation of the CS in the lexicon (Sadock 1998). A phonological account

---

[30] Borer (1996, 1999) claims that the contrast in the wide scope interpretation of cardinals, which is blocked in the CS but available for FS sentences, supports the claim that the syntactic word-properties of the CS are the consequence of incorporation, but Siloni (2003) observes that the availability of wide scope varies greatly with the choice of quantifier and with context, and therefore it cannot follow from the word properties of the CS.

[31] To explain the definiteness spread phenomenon in the CS, the analysis relies on the mechanism of percolation, by which definiteness percolates from N2 to D thereby rendering the whole CS definite, along the same lines argued for in the incorporation account.

[32] This analysis exploits right-peripheral specifiers and right-adjunction, violating Kayne's (1994) antisymmetry hypothesis, which assumes that the basic underlying word order is specifier-head-complement, and all other word orders are derived by movement operations.

(Benmamoun, 2003; Siloni, 2001, 2003) argues that the CS forms a single prosodic domain. Siloni (2001, 2003) further proposes a PF-based analysis of the CS, in which the CS is the prosodic domain for genitive case checking by the maximal projection of N2, thereby deriving the word status of the CS, from which the rest of its properties follow.[33] A semantic account of the definiteness spread phenomenon in the CS (Dobrovie-Sorin, 2000, 2002; Heller, 2002) bases its argument on semantic composition. The definiteness properties of the CS are derived on the assumption that the head of the CS is functional, denoting functions from individuals to individuals; in combination with N2 it derives the uniqueness interpretation of the CS.[34]

In this dissertation, I adopt the head movement analysis of the FS and CS as proposed by Siloni (1996, 1997). Among the many proposals for the analysis of the CS with its unique properties, no one account has yet to address all of the peculiarities presented by this construction. The N-to-D approach in general seems to be a leading proposal in the literature, as it attempts to account for the unique properties presented by the CS while nicely capturing the parallel meaning of the FS and CS by proposing a similar underlying structure (in terms of the base-generated position of the nouns constituting the complex nominal); Siloni's analysis, in particular, is attractive in motivating the underlying structure of the complex nominals, though it is open to some

[33] Siloni (2001, 2003) notes that the head (N1) of the CS shows phonological properties of a functional word (it is a stressless reduced form). Subsequently, she proposes that it qualifies as a function word at prosodic structure and therefore is part of the prosodic word formed by the extended projection of N2. This, in turn, constitutes the prosodic Case-checking domain in which genitive Case is assigned to N2 by the head (N1), which is endowed with genitive Case features.

[34] Both semantic analyses are interesting in that they attempt, contrary to the rest of the proposals in the literature, to provide a principled theory for the interpretation of definiteness spread based on the semantic compositionality theory. They do so by assigning a functional interpretation to N1 in order to derive the uniqueness interpretation (of the type <e>) of the (definite) CS as a whole. (Indefinite CSs are left branching, and have a different argument structure compared to that of definite CSs, according to this proposal.)

objections (see below). Although the discussion below focuses on the head movement approach, I will take advantage of footnotes to provide some context by noting alternatives where relevant.

Siloni's analysis builds on the head movement approach proposed by Ritter (1991), with additional arguments that lead her to somewhat different conclusions, as will be noted below. The analysis of complex nominals is based on examples of derived nouns, specifically those expressing a process or an event, e.g., *harisat ha-cava 'et ha-'ir* (CS); *ha-harisa šel ha-cava 'et ha-'ir* (FS), (both: 'the destruction of the city by the army'); see detailed discussion below. Although the structure of FS and CS is motivated on the basis of evidence afforded by derived nouns, it is generalized to other nouns as well, since they are claimed to exhibit similar properties as those of process nouns.

Based on the observation that the FS and CS are semantically synonymous and interchangeable, Siloni (1997) proposes a uniform analysis of the FS and CS (though their underlying structure is slightly different, as will be shown below). Siloni (1997) interprets Baker's (1988) Uniformity of Theta Assignment Hypothesis in minimalist terms such that it requires thematic paraphrases to have their arguments inserted in the same position. She therefore argues that the position of the arguments must be similar in both FS and CS; in particular, she proposes that in both complex nominals the external argument is generated in specifier position of the noun phrase and the internal argument is generated inside N', as a complement of N (see further details below). Let us consider in more detail first the account for the FS, followed by an account for the CS which addresses the unique properties of this construction.

2.5.1 The Free State

Central to Siloni's analysis of the FS (as well as the CS, see below) is the idea that the thematic representation of derived noun phrases should be similar to that of the verbs from which they are derived. Siloni (1996, 1997) assumes that derived nouns inherit the theta grid of their associated verbs. Thus, in the example in (32) (from Siloni 1996, 1997), the verb *haras* 'destroyed' (32a) and its counterpart noun *harisa* 'destruction' (32b) assign a patient theta-role to *ha-'ir* 'the-city' and an agent theta-role to *ha-cava* 'the-army'.

(32) a. ha-cava haras 'et ha-'ir
the-army destroyed ACC the-city
'The army destroyed the city'

b. ha-harisa šel ha-cava 'et ha-'ir
the-destruction of the-army ACC the-city
'The army's destruction of the city'

According to this proposal, the agent theta-role (*ha-cava* 'the army') receives nominative Case in the verbal construction and genitive Case in the nominal constructions; the recipient of the patient theta-role (*ha-'ir* 'the city') receives accusative case in both verbal and nominal constructions.

The similarity between noun phrases and their verbal counterparts is further explored: Siloni (1996, 1997) assumes that similarly to clauses, Hebrew noun phrases have a hierarchical internal structure in which the patient theta-role (the logical object of the noun, *ha-'ir* 'the city') cannot precede the agent theta-role (i.e., the logical subject of the noun, *ha-cava* 'the army'). For derived nouns, the logical subject must immediately follow the head noun (33a), reminiscent of the word order in Hebrew clauses with post-

verbal subjects, in which the subject must immediately follow the verb (33b) and cannot be preceded by the object (33c); examples taken from Siloni (1996, 1997). She therefore concludes that the logical subject must assume a position higher than that of the logical object.[35]

(33) a. * ha-harisa 'et ha-'ir šel ha-cava
the-destruction ACC the-city of the-army
'The army's destruction of the city'

b. 'etmol haras ha-cava 'et ha-'ir
yesterday destroyed the-army ACC the-city
'Yesterday the army destroyed the city'

c. * 'etmol haras 'et ha-'ir ha-cava
yesterday destroyed ACC the-city the-army
'Yesterday the army destroyed the city'

Siloni presents evidence related to the Binding Theory to support the internal syntactic structure she proposes. She shows that the patient role (e.g., *ha-harisa*) can be an anaphor and as such, according to Principle A of the Binding Theory, it must be bound in its governing category. Since the only antecedent available is the agent, as Siloni notes, then it must c-command the patient. She further shows that the agent role cannot be c-commanded by the patient, and therefore must be structurally more prominent than the patient, and must occupy a higher structural position than the patient or theme (for further details, see Siloni 1996:245-246).

Siloni thus proposes that in the underlying structure of the FS in (32b), the noun phrase that bears the external theta-role, *ha-cava*, is generated in specifier position of the

[35] This proposal is similar to that of Ritter (1986, 1988), who draws a parallelism between clauses and nominal phrases. However, Ritter (1986,1988,1991) compares sentences and process nominals in English: both IP and DP are headed by a functional element (Infl and D respectively), which selects a single complement (VP and NP respectively), as proposed for VSO languages such as Irish and Welsh by Sproat (1985).

NP and the bearer of the internal role *ha-'ir* is generated within the N' beneath that NP. Given the hierarchical internal structure on the one hand, and the surface order of constituents they exhibit on the other hand (e.g., the placement of adjectives, see discussion in §2.2 above), Siloni argues for the raising of the noun to a functional head D, which is motivated by the strong definiteness features that need to be checked and eliminated before spell-out (see further discussion below). The proposed underlying structure of the FS is presented in Figure 2-1 and the resulting surface structure in Figure 2-2. [36] [37]

[36] The analysis of Siloni (1996, 1997) is neutral on the issue of the category of *šel* ('of') and in particular whether it is or it is not a preposition. The DP with the external argument is not further analyzed in terms of its internal structure. The realization of *šel* in the FS may be the result of predicate inversion (den Dikken, 1998), as proposed in Shaked (2001)'s comparison of the Hebrew FS and the 'N of N' construction, based on den Dikken (1998). Den Dikken assumes that inversion of the predicate across its subject gives rise to a copular element; in the example discussed here, the Hebrew *šel*. Copular elements, it is proposed, are not lexical primitives; they are rather the overt realization of functional heads. The spelled out *šel* is meaningless in itself, but it is necessary in order to license predicate inversion. In the underlying structure proposed in Shaked (2001), the possessum (N1) is base generated in specifier position of the small clause, while the predicate of the small clause is a PP headed by a null dative P, of which the possessor is a complement. The surface structure is derived through predicate inversion, which is contingent on head-movement of the functional element in order to extend the domain to which the phrase can move; this, in turn, causes the functional element to be phonetically realized as the copular element *šel*, a linker element.

[37] For FS nominals that are not process nominals (e.g., *ha-bayit šel ha-yeled* 'the house of the boy', Siloni assumes the same underlying structure, in which N1 (*bayit)* is the head N (unlike the structure in Figure 2-1, N does not have an internal argument); N2 (*yeled* 'boy') is the head of DP which occupies the specifier position of NP (the external argument, in Figure 2-1). Thus the structure of non-process FS nominals is identical to that presented in Figures 2-1 and 2-2, but without the final DP (the internal argument).

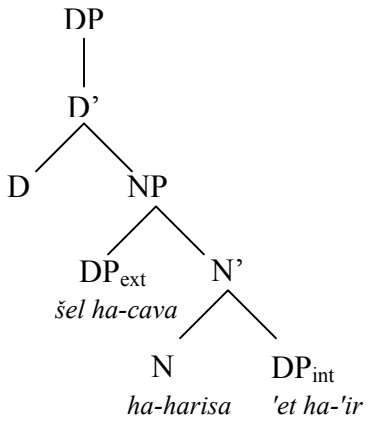


Figure 2-1: FS, underlying structure, from Siloni (1996, 1997)

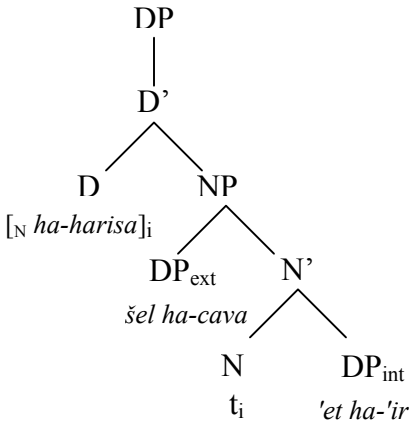


Figure 2-2: FS, surface structure from Siloni (1996, 1997)

The word order of the FS is thus derived through movement of N to D. Siloni (1996, 1997) assumes, following Borer (1996, 1999), that definiteness in Hebrew is a feature on nouns.[38] Hebrew nouns are inserted from the lexicon with definiteness specifications, and these lexical features must be checked with D. The definiteness features in D are strong, and therefore need to be eliminated before spell-out. N1 must therefore raise to D overtly, in order to check the strong definiteness features in D, which surface out in the form of the definite article.[39]

In terms of the placement of adjectives, Siloni assumes that modifying adjectives are base-generated in a position left-adjoined to NP, following Chomsky's (1986) claim that maximal projections adjoin exclusively to maximal projections. After the noun has raised to D, it precedes the adjective, which now appears immediately to the right of the noun, as is normally the case in Hebrew, e.g., *ha-harisa ha-mehira šel ha-cava* 'the rapid destruction of the army'. The position of adjectives raises potential problems in its assumption that adjectives adjoin to the whole NP, and in particular, that the modifier of N1 is adjoined to the maximal projection of N1, above the specifier which harbors the genitive phrase (N2). This may be viewed as problematic for narrow-scope contrastive modifiers of N1 (e.g., *ha-harisa ha-mehira šel ha-cava, lo ha-'itit* 'the rapid destruction of the army, not the slow'), as it would not be possible to semantically focus on N1 alone, to the exclusion of N2. On the other hand, the position of adjectives in Siloni's analysis

[38] Based on evidence for the presence of definiteness concord in adjectival construct states and demonstratives, Borer (1996, 1999) concludes that definiteness is Semitic languages is a feature base generated on nominal heads. Borer further assumes that in Semitic, D is unspecified for definiteness value: it inherits it from a moved N which is specified for definiteness.

[39] Ritter's (1991) and Siloni's (1996, 1997) analyses of the FS differ in that Ritter assumes that N1 raises to Num with no further movement (to D) attested, while Siloni assumes that in the FS, N1 raises directly to D, and no AgrP is projected. This difference between the two analyses stems from the different assumptions related to different features on N1 and to the role of Num vs. Agr with relation to strong definiteness and/or genitive features.

has two clear advantages, as Siloni notes. First, if N1 raising to D did not occur and the noun were to remain in situ, then the analysis would require the ad-hoc stipulation that modifying adjectives are generated in a position right-adjoined to N' in this structure, which is contrary to Chomsky's argument. Second, Siloni's analysis nicely captures the difference between English and Hebrew in the word order of nominal modifiers (in Hebrew, adjectives are postnominal while in English they are prenominal), without the need to assume different positions for adjectives in the two languages. In both Hebrew and English, it is assumed that adjectives are adjoined to the maximal projection NP; the different surface position that adjectives assume in the two languages stems from the need to raise the noun to D in Hebrew, but not in English.

To sum, Siloni's analysis of FS process nominals in particular, and FS nominals in general, crucially assumes a thematically determined hierarchical structure in complex nominals, similarly to that in clauses. According to this analysis, the genitive nominal (N2) is generated in specifier position of NP, which is the maximal projection of the process noun (N1). The surface word order is derived by raising N1 to D, which is required in order to check the strong definiteness features in D that need to be eliminated before spell-out.

2.5.2 The Construct State

Since Siloni (1996, 1997) assumes that the FS and the CS share the same argument structure (see above), the analysis proposes that the CS, like the FS, has the external argument generated in specifier position of NP and the internal argument inside N'. The surface word order is derived by noun raising, similar to that of the FS. However,

the derivation of the CS involves additional movement of the genitive phrase, in order to account for two of its most salient properties: non-local modification by an adjective and non-overt definiteness (definiteness spread).

Siloni's analysis of the CS centers on the mechanism of genitive Case assignment, which she claims also explains adjectival modification in the CS. She presents evidence that the Case on N2 is of a different nature in the FS and CS. In the FS, the genitive Case on N2 is assumed to be inherent Case, assigned by *šel* to the theta-marked external argument (this is argued to be reminiscent of Chomsky's (1986) assignment of inherent Case). When *šel* is present, the external argument receives inherent Case and therefore has no motivation to move. Genitive Case in the CS, on the other hand, is not inherent Case but rather a structural Case assigned under government (as such, according to Chomsky (1986), it is not thematically dependent). The external argument must be raised to a position adjacent to the Case assigner (the raised N1) in order to be licensed.

Siloni follows Chomsky's (1991, 1993) suggestion that structural Case is always checked in Specifier position of AgrP, and proposes that the CS must contain an AgrP and that structural Case in the CS is checked in this agreement projection.[40] Siloni argues, in line with proposals by Ritter (1987, 1988) and Fassi Fehri (1989), that in the CS, the noun is inserted with theta features that it must check with Agr, and correspondingly Agr has N features that it must check with the noun; since the features are strong, they must be eliminated by spell out; therefore N1 raises over its specifier to

[40] It is generally assumed in the Hebrew literature on the CS that Hebrew has abstract genitive Case, an assumption based on a comparison of Hebrew with other Semitic languages. Shlonsky (2004), for example, claims that Hebrew and other Arabic dialects have abstract genitive Case on N2 of the CS similar to the overt genitive Case marking in Standard Arabic, see example from standard Arabic CS in (i).

(i) dar-u r-rajul-i
apartment-NOM the-man-GEN
'the apartment of the man'

the functional head Agr and incorporates with abstract Agr. Agr, once supported morphologically, assigns genitive Case to the specifier position of AgrP. The genitive DP (the external argument), base generated in spec;NP1, raises to spec;AgrP, in order to be assigned genitive Case (the DP features of Agr are also strong, therefore must be eliminated before spell-out).[41] [42] [43] This movement is followed by subsequent movement of the functional head (Agr) containing N1 to D, since D also has strong N features that need to be checked (as noted above). Thus, Siloni (1997) proposes three movement operations in the derivation of the CS: N → Agr_{gen}; DP_{ext} → SpecAgr_{gen}P; and Agr_{gen} → D. The underlying structure and derivation of the CS, and its resulting surface structure, are presented in Figures 2-3 and 2-4, respectively.[44]

---

[41] Ritter (1991), Siloni (1996, 1997) and Longobardi (2001) speculate on the morphophonological alternations of the head of the CS. Ritter (1991) claims that they result from the genitive feature on D; the N1 bound form serves to make the genitive feature visible. Siloni (1996, 1997) notes that it is a reflex of the fact that Agr has strong features. Longobardi (2001) attributes it to the combination of the lexical entry of N1 with an exclusive prosodic (non-segmental) morpheme in D, which might suggest the effect of an abstract phonological trigger in instances of N-to-D raising. See §2.3 above for a phonological account of the morphophonological alternations from treatment of the CS nominal as a prosodic word and consequently the domain for stress assignment rules (resulting in stress reduction of N1).

[42] Longobardi (1996, 2001) proposes an alternative trigger for the head movement according to which the strong definiteness feature in D motivates head movement. He claims that N-to-D raising is a cross-linguistic phenomenon that applies in all DPs, and that languages vary in whether the raising is overt or covert, depending on the feature in D. He argues that it is crucial to keep the motivations for fronting of the head N1 conceptually disjoint from the process of genitive checking: the N-to-D movement (via Agr) is functionally motivated by the need to check the article feature, while genitive Case assignment is the consequence of dragging to D an Agr which is in a spec-head relation with a genitive argument (N2). This analysis is designed to capture the relation between raising N to D and the assignment of genitive Case, without triggering the movement by Case checking.

[43] Both Siloni (1996, 1997) and Ritter (1991) motivate movement of the extended projection of N2 by the need for it to be assigned Case; however, Ritter (1991) assumes that N2 is assigned Case by D (via spec-head agreement between D and the specifier of NumP), while Siloni (1996,1997) assumes Case is assigned by Agr in a spec-head configuration between the head (Agr) and its specifier.

[44] CS nominals that are not process nominals (e.g., *beyt ha-yeled* 'the house of the boy') are assumed to have a similar underlying structure (see footnote 37 for the underlying structure of non-process FSs): N1 (*beyt)* is the head N and N2 (*yeled* 'boy') is the head of the DP in specifier position of NP (the external argument, in Figure 2-3). Thus the structure of non-process CS nominals is identical to that presented in Figures 2-3 and 2-4, but without the final DP.

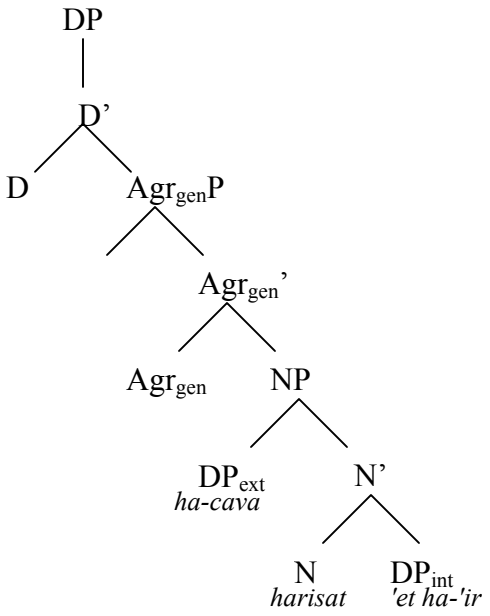


Figure 2-3: CS, underlying structure, from Siloni (1996, 1997)

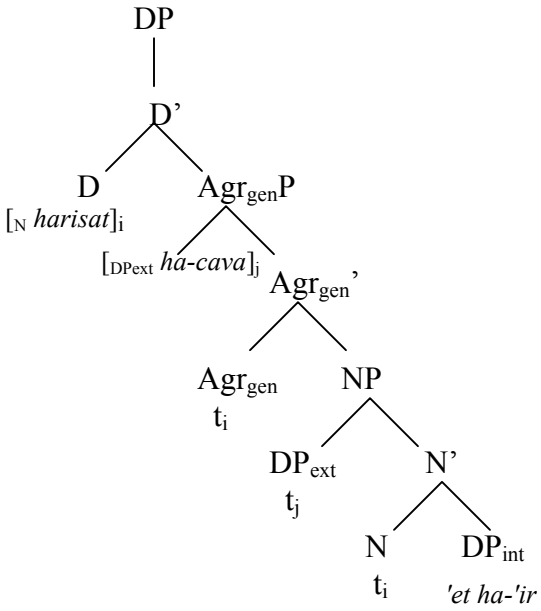


Figure 2-4: CS, surface structure, from Siloni (1996, 1997)

Siloni (1997) motivates the presence of an AgrP in the CS by the need for its specifier to be the locus for Case checking for the genitive (external) argument. The position of modifying adjectives in the CS supports this structure. Recall from the discussion in §2.2 above that an adjective modifying the head (N1) of the CS cannot immediately follow N1, a restriction entailed by the adjacency restriction. Consequently, adjectival modification of the two nouns follows the CS in a nested configuration. Based on the linear word order of adjectives in the CS, Siloni shows that the position to which the genitive phrase moves must precede adjectives modifying the head noun. Similarly to the position of adjectives in the FS, Siloni assumes that adjectives in the CS are base-generated in a position left-adjoined to NP. Assuming that a functional projection (AgrP) above the NP and above its left-adjoining adjectives is necessary in order for the Case assigner (the head N) and Case assignee (the genitive phrase) to be strictly adjacent for Case assignment to be licensed, an intervening adjective between the head and the genitive phrase would block case assignment. Therefore the genitive phrase (headed by N2) must raise to spec;AgrP, where it checks the strong genitive features in a spec-head configuration, as discussed above. Note that Siloni assumes a similar position for adjectives in both CS and FS; the differences between their surface structures stems from the assumption of an AgrP above the NP projection in the CS (but not in the FS) to which the genitive phrase overtly moves, thus deriving the nested configuration in the CS.

The presence of the Agr projection also explains the definiteness peculiarities of the CS, according to Siloni. As shown in §2.2, the head noun (N1) in the CS cannot have a definite article prefixed to it. In addition, the two nouns in the CS must agree in their definiteness value, as is evidenced by the presence of the accusative marker *'et* preceding

a definite CS and by the definite article prefixed to adjectives modifying a definite CS (see §2.2 for examples). Genitive relations in which the two nouns do not share their definiteness value must be expressed through the FS, not the CS. Siloni compared the absence of a definite marker on N1 of the CS to the constraint on the co-occurrence of possessors and articles in the English Saxon genitive (e.g., **John's the book*). This complementary distribution of possessors and articles in the English Saxon genitive is explained by Abney (1987)'s 'article constraint': the possessor *'s* is in spec;DP and the article *the* in D cannot co-occur. In Hebrew, Siloni claims, a similar constraint applies. She assumes that Agr and the definite article are incompatible because both include a definiteness feature, which cannot have two realizations on the same head noun. Siloni proposes that in Hebrew, a noun can realize its definiteness features either as part of its agreement morpheme or in the customary way, in the form of an article. If a noun is inserted with genitive Agr features, its definiteness feature is part of the Agr features, and the article cannot appear. Genitive Case is assigned structurally by Agr, as argued above. If the noun is inserted with the article, it cannot bear Agr features, therefore an AgrP is not projected and the noun raises immediately to D, as is the case in the FS. Thus N1 can either be inserted with Agr features, which give rise to a CS, or with the definite article, which gives rise to an FS, but not with both.

Siloni makes another assumption, in order to account for the definiteness agreement between the two nouns in the CS. She assumes that when the definiteness features of the noun are part of Agr, although the noun is not specified for definiteness in the usual way (with an article), it does not remain underspecified; rather, the noun (in D) must agree with the definiteness value of its complement, i.e., it must match the

definiteness value of the genitive DP.[45] This explains the definiteness agreement among members of the CS as a feature matching or feature agreement between the definiteness feature in of Agr (in D) and that of DP2 (in spec;Agr).[46] It should be noted that Siloni's (1996,1997) assumption that definiteness spread is the result of definiteness matching or agreement between N1 and N2 (a result of spec-head agreement between D and spec;AgrP (the position to which the extended projection of N2 raises) faces some objections.[47] Nonetheless, no other proposal in the literature on definiteness spread in Hebrew seems to better explain the definiteness facts of the CS.

To summarize: the syntactic analyses proposed by Siloni (1996, 1997) for the derivation of the FS and CS are interesting in that they attempt to capture several important differences between the two constructions, namely the lack of a definite article on N1 of the CS, and the nested configuration of modifying adjectives. The account takes as its starting point the fact that the FS and CS, though different in their surface structure, are derived from a similar underlying argument structure and must be so since they are essentially paraphrases of each other; that the CS and FS are paraphrases of each other is a guiding assumption in this study of the processing of complex nominals in Hebrew.

---

[45] Siloni (1996) mentions cases of agreement between a Case assigner and its assignee in a structural configuration of government to support this assumption (for example, a verb governing its postverbal subject in Hebrew shows agreement with it). She also notes that agreement in definiteness value is not unique to the CS; adjectives and nouns they modify also agree in their definiteness value in Semitic, thus supporting the claim that (in)definiteness is a feature in which constituents can agree.

[46] See Danon (2008) for an alternative minimalist explanation of definiteness spread in the CS in terms of the operation Agree as a feature sharing operation passing feature values from one node to the other, at the syntax-semantic interface. Danon also gives a detailed and critical review of the different accounts for definiteness spread proposed in the vast literature on the Hebrew CS, some of which have been touched on briefly above.

[47] One objection might be that the assumption of semantic sharing as a result of spec-head agreement results in the wrong generalization for the spec-head agreement mechanism, which is often used for structural feature checking. Furthermore, while it has been attested that one feature is multiply checked (for example, gender features), it is unattested that a single feature has several realizations. It is widely assumed that a definiteness feature can have only one realization as a definite article.

2.6 Summary

This review of the literature on the linguistic analysis of Hebrew complex nominals leads to the conclusion that the FS and CS are very different in their phonological and syntactic propoerties, yet they are used to convey identical meaning. Phonologically, it has been shown that the FS consists of two independent prosodic words, each carrying its own phonological stress, while the CS exhibits properties that have been shown to follow from its status as a single phonological word. Syntactically, Siloni argues that the FS and CS are derived from a similar (though not identical) underlying argument structure, which needs to be assumed in order to account for the semantic similarity between them. The only difference in the deep structure between the two constructions is the additional abstract functional category AgrP in the CS structure, which needs to be present for the derivation of the CS to be licit, according to this analysis. This necessitates two steps of movement for the head noun, since it must pass through Agr. Also, the derivations of the two complex nominals are different since in the FS only head movement is required while in the CS there is also raising of the genitive phrase headed by N2. Thus the resulting surface structure of the CS differs from that of the FS not only in that it has an additional functional head, but in that its genitive phrase is higher in the tree structure than that of the FS, which has not moved during the course of the derivation and remains beneath N'. These differences might be invoked to explain processing differences that are observed between the two constructions; see discussion in Chapter 8. What seems clear, however, is that semantics could not be the source of any processing differences between them. Although the FS and CS constructions have different properties with respect to both their syntactic structure and their phonological

structure, it has been argued that the FS and CS do not differ in their semantic interpretation. It was shown that a wide range of relations may be conveyed by both the CS and the FS, but it was argued that any semantic differences between them are only a matter of usage, with the FS more frequent in spoken language, particularly in colloquial speech, while the CS is more frequent in written language, and in formal speech. Both, however, are perfectly familiar to native speakers of Modern Hebrew and are often used interchangeably.

CHAPTER 3

HEBREW PROSODY: AN OVERVIEW

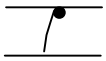
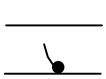
The literature on Hebrew prosody is scarce. In the words of a notable scholar, Isre'el (2005), "Whereas phonetics and phonology of Israeli Hebrew […] have been dealt with in a number of significant studies, work on prosody has been extremely scanty, and can be ascribed mainly to a single scholar, viz., Laufer (1987)." (p. 5). In this chapter I review Laufer's comprehensive work, which includes an inventory of nuclear tones and intonation patterns, and the cues for identifying prosodic phrase boundaries. A few other works that are relevant for this discussion will also be mentioned. These facts will contribute to the discussion of Hebrew prosody in subsequent chapters. In recent years, related work on the properties of Hebrew prosody in spontaneous speech has been a focus of study (Amir, Silber-Varod & Isre'el, 2004; Silber-Varod, 2005); I report these works and discuss their relevance. Finally, a characterization of stress assignment in the Hebrew nominal system will follow; as suggested in Chapter 2, §2.3, the properties of the CS follow from its status as a prosodic word; therefore, an investigation of the literature of stress assignment in the nominal system provides a general background for the characteristics of stress assignment in the CS.

3.1 Inventory of nuclear tones and intonation contours in Hebrew

Laufer (1987) identifies 6 main tone levels (nuclear tones) in Hebrew. The inventory of these tones consists of single accents, and accent-accent sequences, as presented in Table 3-1. An adaptation of Laufer's level-based approach into the current

standard Autosegmental-Metrical (AM) approach (Ladd, 1996; Pierrehumbert, 1980) is outside the scope of this dissertation, therefore the discussion below maintains the terms used by Laufer.[48] Each tone in the table below is described and presented with the schematic contour used in Laufer (1987), which appears between two horizontal lines representing the frequency range.

Table 3-1: Inventory of Hebrew nuclear tones, as presented in Laufer (1987) in Hebrew (pp.32-33) and translated into English by AS. The symbols are to be read from right to left (following the directionality in Hebrew orthography).

	Nuclear tone	Schematic contour	Description
1	A small falling tone		The tone falls throughout the word from middle to low tone level.
2	A large falling tone		The tone falls throughout the word from high to low tone level.[49]
3	A rising-falling tone		First the tone rises from a low to a high tone, and then it quickly falls to a low tone level.[50]
4	A small rising tone		The tone rises throughout the word from a low tone level to a medium (or slightly higher) tone.[51]
5	A large rising tone		The tone rises throughout the word from medium (or lower) tone to a high tone.[52]
6	A falling-rising tone		The tone falls from a relatively high to a low tone, and continues to rise to a medium tone.

[48] The Autosegmantal-Metrical approach to intonation (a term coined by Ladd, 1996, who developed the approach from the work of Pierrehumbert, 1980) uses usually two (High and Low) and sometimes three (High, Low and Middle) levels for describing intonation, quite similar to Laufer's characterizations of the tone levels in Hebrew.

The first three tone levels are grouped into what Laufer calls ‘falling tones’, while the last three tones are ‘rising tones’. In intonation contours with small or large falling tones (1, 2), the falling tone occurs on the first (or sometimes second) syllable of a phrase, while the remaining syllables up to the end of the phrase remain in a relatively low tone level. A rising-falling tone (3) stretches over the two or three first syllables: the tone is rising on the first (or sometimes second) syllable, and falls in the second (or third) syllable to the lowest tone level. The remaining syllables of the phrase remain more or less in that low tone. In intonations with a small or a large rising tone (4, 5), the tone rises throughout the sequence of syllables: the first syllable has a low tone, and it slides upward with each consecutive syllable (each syllable is higher in tone than the preceding one), until it reaches its peak at the end of the last syllable. The final tone level is middle in a small rising tone and high in a large rising tone. In a falling-rising tone (6), the falling tone is on the stressed syllable of the accented word, and the rise in tone is on the last syllables of the phrase, while the middle syllables have a low tone.

Laufer claims that every intonation pattern in an intonational phrase must end with one of the six nuclear tones above. Every phrase has a nucleus, which is the stressed syllable of the last accented word, and it may be preceded by a ‘head’ (which is the stressed syllable of the first word in the prosodic phrase, up to the nuclear tone). The head

---

[49] Laufer notes that acoustic analyses show that a large fall is preceded by a rise, although hearers perceive it as a large fall, because in order to produce a large fall, speakers usually start to speak at a tone in the middle part of their range and need to reach a tone at the top of their range.

[50] Laufer claims that hearers differentiate between a rise preceding a large fall (see footnote 49) and a rising-falling tone, since the rise in the latter is more substantial (and always occurs), it lasts longer and starts at a lower tone.

[51] As Laufer notes, acoustically, a small rise can be realized even as a horizontal tone (without a rise or a fall). He explains it as a natural tendency towards a lowering of the fundamental frequency during speech. This tendency has the consequence of lowering the rises and magnifying the falls.

[52] A large rise, especially when the syllable onset is voiced, can be realized acoustically as preceded by a fall, for the same type of reason that a large rise is often realized acoustically as preceded by a fall (see footnote 49).

may be preceded by pre-head tone(s) (the syllables preceding the stressed syllable of the head). Laufer lists two types of pre-heads and two types of heads. In addition to a head and pre-head(s) preceding the nuclear tone, the nuclear tone may be followed by a tail. Laufer also provides a description of accentuation of a specific word within the phrase (for focalization), which may be achieved by changing the frequency range, or changing the form of the head.

Laufer provides a detailed account of 9 common sentence intonation patterns in Hebrew (for the contexts in which they occur and/or emotional meaning which they impart with respect to sentence modality, see a detailed discussion in Laufer's work). In general, Laufer notes that the intonation contours ending with a falling nuclear tone express completion and finality while rising nuclear tones express incompleteness and continuation. Table 3-2 lists the 9 sentence intonation patterns described by Laufer (and translated by AS), as well as a short description of each contour (see Appendix 4 in Laufer, 1987 for the schematic contour Laufer provides for each).

Table 3-2: Sentence intonation contours in Hebrew, translated from Laufer (1987), with a short description of each contour.

	Intonation contour	Description
1	(low pre-head+)(low head+) small fall (+tail)	A falling contour throughout the sentence, ending in a small falling nuclear tone.
2	(low pre-head+) high head + small fall (+tail) or: high pre-head + small fall (+tail)	A low contour which rises (or a high contour), ending with a small falling nuclear tone.
3	(low pre-head+)(high head+) large fall (+tail)	A rising contour throughout ending with a large falling nuclear tone.
4	(low pre-head+) low head + large fall (+tail)	A low contour throughout the sentence ending with a large falling nuclear tone.
5	(low pre-head+) (high head+) rise-fall (+tail)	A low then high contour ending with a rising-falling nuclear tone.
6	(low pre-head+) (low head+) small rise (+tail)	A low contour ending with a small rising nuclear tone.
7	(low pre-head+) high head + small rise (+tail) or: high pre-head + small rise +tail)	A low contour then high, with a small rise in nuclear tone.
8	(low pre-head+) (high head+) large rise (+tail)	A low followed by high contour, ending with a large rising nuclear tone.
9	(low pre-head+) (high head+) rise-fall (+tail)	A low, then high contour, ending with rising falling nuclear tone.

As Laufer notes, in general, a falling intonation is usually found at the end of a declarative sentence, while a big rise marks a yes/no question. A rising intonation (small rise/fall-rise) often marks an expectation of continuation of the sentence, or hints that

while the information in the 'sense group' (an intermediate phrase, see §3.2 below; this characterization will serve as basis for the analysis presented in Chapter 5) is not new/interesting, the following unit should be of interest. The continuation intonation (small rise and fall-rise) is claimed to be characteristic of complex sentences.

In a preliminary characterization of the basic intonation patterns in Hebrew,[53] Chayen (1973) provides similar characterizations of the nuclear tones associated with different modalities: in declarative sentences, "the nuclear tone of the final breath group […] is a falling tone and indicates finality of the utterance." (p. 35) A falling tone is also associated with wh-questions, commands and exclamations. Yes/no questions have a rising tone on their last stressed syllable. With respect to non-final breath groups, he notes that a rising nuclear tone is used and adds that non-final rising tones signaling a continuation may occur at the end of a subordinate clause preceding a main clause or vice versa, or in slower speech when the utterance is broken up into shorter units. A rising tone is often replaced by a falling-rising tone (and considered a variation of the plain rising tone), particularly on the nuclear syllable of a breath group which is non-final in the utterance. As speech becomes more rapid, he notes, the phrase becomes longer as well, and there are fewer (or weaker) stresses within it.

Effects of focus in Hebrew are also of interest, and will be further discussed in Chapter 7. Mixdorff & Amir (2002) investigate intonation patterns with relation to focus. In their study, the simple sentence *amarti lahem lavo* 'I told them to come' was uttered by one speaker in 8 different contexts (each uttered 5 times), providing 40 target

[53] Chayen (1973) notes that "a comprehensive description of Hebrew intonation patterns reflecting all the attitudes and emotions which may accompany utterances would require a larger and more varied set of recordings…", since the preliminary characterization is not detailed and is based on "'cold', unemotional, unemphatic monologues" such as newscasts, radio broadcasts, a press conference etc. (p. 34).

sentences for their analysis. The contexts included broad focus and narrow focus on different words of a declarative sentence, echo questions with broad and narrow focus, and neutral non-terminal contexts (where the simple sentence is the matrix clause of a complex sentence consisting of either two or three clauses, e.g., *Amarti lahem lavo aval hem lo racu ki hem racu lalexet lemakom axer* 'I told them to come but they didn't want to because they wanted to go to a different place'). Despite the admittedly small amount of data examined in this preliminary study by Mixdorff & Amir (2002), the quantitative analysis found that words that have narrow focus have increased durations relative to their occurrence in sentences with broad focus.

With regard to sentence mode, Mixdorff & Amir found that declarative sentences are characterized by a falling tone on the final syllable. A question intonation is marked by a steep rise on the final stressed syllable (of more than 1.5 octaves), and was accompanied by a lengthened final syllable. Rises in the final syllable and final syllable lengthening were found at the end of the basic phrase, when it was embedded in a complex sentence and followed by another phrase (see above). An effect of "speeding up" was found when the sentence uttered was longer (a complex sentence): a comparison of measurements of the first 5 syllables of the main clause showed that when the clause was a simple sentence, the first 5 syllables were 146 milliseconds (ms) in duration on average while in contexts in which it was part of a 2 clause sentence, they were 135 ms in duration. To summarize briefly: Mixdorff & Amir (2002) found the following intonational patterns in Hebrew: Falls on declarative sentences, rises in non-terminal

positions, and very strong rises for question-final accents.[54] These findings correspond to Laufer's (1987) perceived intonation patterns, as described above.

Amir, Almogi & Gal (2004) investigated the perception of prominence in short sentences (e.g., *ha-yeled kana lexem* 'the boy bought bread'), which varied in terms of narrow focus, and in levels of final rise.[55] Their study shows that native speakers of Hebrew judged significantly more utterances as having narrow focus, as compared to non-native speakers. In addition, native speakers exhibited categorical perception, when a pitch rise occurred between 20 and 30 Hertz (Hz), while non-native speakers did not do so. The authors conclude that this is the value of pitch increase that is necessary in order to create the perception of prominence. A pilot study reported in their paper showed that subjects can identify smaller changes in pitch (less than 10 Hz), but need a larger change in order for it to be construed as conveying prominence.

3.2 Phrase boundaries in Hebrew

Laufer (1987, 1996) also (briefly) describes levels of what can be translated as "prosodic disjunction" in Hebrew. He notes that it is possible to order levels of disjunction between words on a scale, according to the loose or tight semantic connection between groups of words (i.e., a very loose connection exists between groups of words that concern different topics; a tighter connection may hold between the subject and the

[54] In cases where the last syllable in a question was unaccented, a boundary tone was observed on the last syllable.

[55] The study compared two groups of participants: Native Hebrew speakers, and Native Arabic speakers who are L2 Hebrew speakers at a high academic level. Each sentence was played with 20 variants, differing in intonation contour by manipulating (1) Narrow focus (there were 5 levels of increase on the stressed syllable, from 10 to 50 Hz) and (2) Inquiry (there were 5 levels of final rise, from 20 to 100 Hz). Participants' task was to choose from 5 responses to indicate whether the sentence conveyed no additional meaning, or whether an additional meaning was expressed by the utterance of the sentence (4 specific meanings were presented for each sentence, e.g., for the sentence 'the boy ate bread', one option would be 'boy, not girl, bought bread').

predicate within a sentence; an even tighter connection holds between a head and its complements/adjuncts). However, he proposes following Jones (1956, 1964) merely a dual division between types of phrases in Hebrew: A 'breath group' and a 'sense group'.

A 'breath group' boundary (marked with two vertical lines (||) and equated with Pike's (1958) 'final pause') is characterized as the boundary between two breaths, and usually corresponds to the end of the sentence. The boundary following a breath group is characterized as longer than the break after a sense group, and usually includes a breath. The characterization of the breath group seems to correspond quite closely to Beckman & Pierrehumbert's (1986) Intonation Phrase (IP).

A 'sense group' boundary (marked by a single horizontal line (|), as in Pike's (1958) 'tentative pause') is a secondary division of the breath group; a breath group may be divided into smaller groups (sense groups), usually in order to convey smaller units like a phrase or even a single word. There is a close syntactic relation between successive sense groups. Even if the break after a sense group is long in duration, it does not include a new breath. A sense group can include a single word, or several words. Its characterization corresponds quite closely in my judgment to an intermediate phrase (ip) in Beckman & Pierrehumbert's (1986) work. Notice that Laufer seems to mix phonetic and semantic criteria in his characterization of prosodic disjunction in Hebrew.

Laufer (1987) notes that several sense groups (assumed here to correspond to ips) form a breath group (assumed here to be an IP) and thus, the end of the last sense group (ip) coincides with the end of the breath group (IP): The breath group (IP) is characterized by a gradual narrowing of the tone range, from a wide range in the first sense group (ip) to a smaller range in the following sense group (ip). In every sense group

(ip), the high tone is slightly lower than the high tone in the previous one. In the following IP there is resetting, and the range of tone is wider again. Laufer also notes, when speaking of sense groups, that "it is difficult to point out precise criteria for the division of the speech sequence into word groups. Nonetheless, it seems that the division is determined by criteria such as syntax and duration." (Laufer 1987: 20; translation by AS). On the one hand there is a tendency to place a prosodic boundary at syntactic edges; on the other hand, there seems to be a limit on the duration of a phrase, which also triggers the division of speech into smaller units.

Three cues are associated with an intermediate phrase boundary in Hebrew (Laufer, 1987, 1996; see also Enoch & Kaplan, 1969 for a description of these cues in Hebrew): a changing tone (pitch height), a changing pace (lengthening), and a break (a physical pause):

1. A change in pitch height. Two tonal events occur at the end of an ip: A change in pitch from the last accented syllable (the nucleus tone) to the ip boundary, and re-setting of the pitch after the ip boundary. The two cues complete and reinforce each other: before the ip boundary there is a change in the nucleus tone, and after the boundary there is re-setting of the pitch height. This was described in Laufer (1987) as the most reliable cue for an ip boundary.
2. Lengthening of the final syllable of the ip, and fast speech rate at the beginning of the following ip boundary.
3. A physical pause. This cue was deemed to be an unreliable indicator of a boundary, which rarely occurs alone and is usually accompanied by other cues, since the end of an ip is not always signaled by a pause.

Laufer (1987, 1996) and Enoch & Kaplan (1969) both argue that the most important cue for an ip boundary in Hebrew is a change in the pitch height. This cue is always present in an ip boundary, but it is often accompanied by the two other cues. Lengthening seems to be a common cue as well, though sometimes absent altogether, while a physical pause is only a secondary cue, not reliably present. This is supported by Laufer (1996), which reports an experiment of the perception of ip boundaries in speech taken from television interviews. Participants identified ip boundaries, and the duration of the identified phrases and pauses was then measured. The results show that perceivers did not associate a boundary solely with the absence of a pause, as an instrumental analysis indicated. The 158 ips which were perceived had breaks which varied in duration from 0 to 117ms, and averaged at 22.08 ms. In 40% of the cases in which an ip boundary was perceived, there duration measurements did not indicate that there was a pause in the acoustic signal.[56]

In a study of spontaneous discourse, Amir, Silber-Varod & Izre'el (2004) and Silber-Varod (2005) investigated the frequency of occurrence of four prosodic cues which can signal the presence of prosodic boundaries in naturally occurring spontaneous speech[57] (as opposed to Laufer's ranking of prosodic cues based mostly on planned/elicited utterances). The working assumption in these studies is that the basic unit of spontaneous language is the prosodic unit (which corresponds to Halliday's (1989, 1994) information unit), disjoint from the structural-semantic characterization of the

[56] These are average figures. The study found gender differences in participants' interpretation of breaks: men were less likely to associate a break with a perception of a pause, as compared with women: 52% of perceived boundaries for men did not have any pauses in the duration measurements, as opposed to women, whose perception of breaks usually coincided with a break (only 15% of perceived boundaries did not have a pause). The study reported by Laufer is from a 1993 term paper by Ms. Gila Drori for the course: Hebrew Intonation.

[57] The corpus includes narratives taken from free conversations: Amir, Silber-Varod & Izre'el's (2004) preliminary study was based on a short narrative from one speaker engaging in conversation with the experimenter. Silber-Varod (2005) analyzed 3 long narratives.

unit.[58] Therefore, the 'Intonation Unit' (IU) they refer to includes (but is not limited to) sentences, clauses, fragments (some ending with hesitations) and false starts, and no distinction is drawn between them (see also a corpus analysis of the types of sequences under the rubric of IU in Izre'el, 2005)[59]. The four cues discussed with relation to an IU are (1) a pause; (2) final syllable lengthening or slow speech rate at the end of an IU, and (3) a fast speech rate following the end of an IU; (4) pitch reset. Thus it appears that the IU is defined by assuming a boundary where some or all of the acoustic cues occur, regardless of the syntactic form of the resulting unit. It seems inadequate to compare the unit described in these studies and the intermediate or intonational phrases mentioned thus far; however, the prosodic cues seem to partially resemble those identified by Laufer (1987, 1996) and Enoch & Kaplan (1969) for an intermediate phrase, and so are worth mentioning here; due to the small number of studies on Hebrew prosody, they may provide additional information in the form of quantitative measurements of prosodic cues in Hebrew.

The study reported in Amir, Silber-Varod & Izre'el (2004) found that the average number of written words in the unit they investigated (the IU) is 3, ranging in length from 1-6 words. The average number of syllables in an IU is 6 (ranging from 1-14 syllables). The average duration of an IU is .95 seconds (ranging from .31-1.78 seconds). In terms of presence of prosodic cues, only 24% of the IUs had all 4 prosodic cues: tone rise, lengthening, pause and acceleration.

---

[58] This is a departure from the assumption of Pierrehumbert (1990) and Selkirk (1980, 1984), adopted in this dissertation, that prosodic phrasing is a function of the syntactic and/or semantic structure as well as other purely phonological factors.

[59] Izre'el's (2005) study of IUs is part of a preparatory stage towards the compilation of the Corpus of Spoken Israeli Hebrew (CoSIH). The corpus includes a variety of texts spoken by diverse people and from a variety of contextual situations.

In contrast with the ranking of cues for a sense group (ip) in Laufer (1987) and Enoch & Kaplan (1969), Amir et al. (2004) and Silber-Varod (2005) present a different ranking of the prosodic cues in terms of their prevalence.[60] Final syllable lengthening was the most prevalent cue for a prosodic boundary (81% of units).[61] Pitch reset (where the difference between the stressed and unstressed syllable was greater that 15 Hz) was also often a cue for a boundary, though it was less common than lengthening (it appeared in 70% of IUs). Importantly, in the majority of the cases the reset was downwards (the average pitch reset was 51 Hz). This supports the assessment mentioned above that the unit studied by this research group is different than the intermediate phrase unit, since the latter is characterized by a rising or a falling rising pitch (see above). Pauses (measured as at least .02 seconds in duration) were found to be present in roughly half of the IUs studied. A fast speech rate at the beginning of the IU following a boundary was found to be the least relevant cue in their study, and never appeared as the only cue.[62]

The difference in prevalence of prosodic cues in this study compared with prior studies is attributed in Amir et al. (2004) to the different sources of data (spontaneous vs. elicited speech). It is also possible to interpret the difference as stemming from the nature of the unit under investigation. Laufer (1987) lists these cues as indicative of a sense group (an ip) which is a subdivision of the sentence into smaller phrases, while Amir et al.'s (2004) IU may range from a false start to a whole sentence. It seems that for the

---

[60] The rankings of Amir et al. (2004) and Varod-Silber (2005) are similar; the data reported here are from Amir el al.'s study.

[61] A final syllable was considered lengthened when the duration of the terminating syllable was greater than 1.1 seconds (this cutoff is 10% more than the average of syllable duration). Average lengthened syllable measured .24 seconds.

[62] For units longer than 4 syllables only, the ratio between the average of syllable duration preceding the prominent syllable and the average syllable duration in the entire unit was measured in order to determine the speech rate. When the rate was less than .9 seconds (syllable duration was 10% less than the average), it was considered a fast speech rate. The average rate was .76.

purposes of distinguishing intermediate phrase boundaries in elicited productions, which is part of the experimental component of this dissertation, the studies of Laufer (1987) and Enoch & Kaplan (1969) are more relevant.

Finally, Koriat, Greenberg & Kreiner (2002) observe that longer pauses in sentences occur at the major syntactic phrase boundaries.[63] They illustrate this with a sentence like (1) below.[64] They measured the duration of silent pauses in predetermined syntactic phrase boundaries, where pauses were expected to occur, and noted that the longest pauses occurred between the NP and the VP, and between the main clause and the relative clause. (The slash sign (/) marks the predetermined syntactic phrase boundaries measured in the reported study; the translation in (1b) is from their article – the original sentence in Hebrew (reconstructed by AS in (1a) based on the English translation) is not presented in their paper.)

(1) a. ha-xatul ha-šamen/$_1$ 'im ha-pasim ha-'aforim/$_2$ rac/$_3$ maher/$_4$ la-xatul ha-katan/$_5$
the-cat the-fat with the-stripes the-gray ran quickly to.the-cat the-small

še-'ibed/$_6$ 'et darko/$_7$ ba-rexov ha-so'en
that-lost ACC way.its in.the-street the-noisy

b. The fat cat /$_1$ with the gray stripes /$_2$ ran /$_3$ quickly /$_4$ to the little kitten /$_5$ that lost /$_6$ its way /$_7$ in the noisy street.

[63] Koriat, Greenberg & Kreiner's (2002) research is aimed at evaluating their hypothesis that in sentence processing during reading, syntactic structure is extracted prior to semantic interpretation; they claim that prosodic structure (both reading aloud and silently), demonstrate that. They show that there are similarities in prosodic patterns in sentences with similar structures (even when sentence structures are lexicalized with nonce-words, as long as the structure is recognizable), but when syntactic structure is degraded (resulting in ungrammatical sentences), prosodic similarities greatly decrease.

[64] The sentence in (1) is part of an experiment they conducted in order to examine prosodic pause patterns of different sentence types read aloud by 6 female students who are native Hebrew speakers. Participants were asked to read the sentence as soon as it was presented (i.e., on a first pass through the sentence).

While most pauses ranged between roughly 40 and 250 ms, the pauses between NP and VP (/2), and between the main clause and the relative clause (/5), were between 500 and 550 ms (Koriat et al. 2002:277).[65] While it is not a novel finding that prosodic patterns and syntactic structures are aligned[66], it is nevertheless relevant to note here that when a relative clause unambiguously modifies a single host noun (in this case, 'kitten'), there is a prosodic break before the relative clause. This boundary preceding a relative clause is particularly interesting with respect to my study of ambiguously attached relative clauses. It is consistent with the view that a long relative clause in Hebrew induces a prosodic break before it, as will be argued in subsequent chapters. Also of relevance, Koriat et al. note that a pilot study indicated that in normal reading, participants' prosodic phrases are 2-3 words in length. This suggests rather short prosodic phrases in Hebrew, as is also confirmed in the production study reported in this dissertation (Chapter 5).

3.3 Stress assignment in the Hebrew nominal system

Chapter 2 argued for the status of the CS as a prosodic word (see, in particular, §2.3): the main stress falls on the second noun, while the stress on the first noun is reduced to a secondary stress. It was argued that this is similar to stress assignment in Hebrew nouns in general; this section reviews further the literature on stress assignment in Hebrew nominals. It seems to be a consensus in the literature that the Hebrew metrical system is a lexical accent system, in which the default stress is assigned to the rightmost accented syllable (word-final stress).

---

[65] Koriat et al. do not report the exact values, but from the chart they present it is possible to estimate the pause-lengths in milliseconds for the sentence above: 250/1, 550/2, 0/3, 190/4, 500/5, 40/6, 150/7.

[66] It should be noted that the lengths of the constituents may have contributed to Koriat et al.'s findings. The break pattern might be quite different for the sentence 'the cat ran quickly and quite fearlessly to the kitten that cried'.

It has been generally observed (Bat-El, 1993; Bolozky, 2000; Chayen, 1973; Graf, 2000; Ussishkin, 2000; among other) that while the main stress is assigned as a default to the final syllable, secondary stress is assigned to every other syllable to the left of the main stressed syllable. Thus, secondary stress is determined according to the location of the main stress. However, this phonological demand of the Hebrew system conflicts with the fact that stress assignment in Hebrew nominals depends on lexical accent specifications on stems and suffixes (inflectional suffixes, indicating gender and number). Thus, the metrical system of Hebrew displays an interaction between a default rhythmical pattern and lexically specified requirements on the nouns. While the phonological rule assigns stress to the final syllable (e.g., the (a) examples in (2) and (3) below), lexically specified stems may require stress placement on them (2b), and suffixes may require marking of stress on the syllable preceding them (3b).

(2)	a.	gamád dwarf	gamadím dwarfs
	b.	tíras corn	tírasim corn cobs
(3)	a.	arnáv rabbit	arnévet rabbits
	b.	yéled child	yeladím children

Graf & Ussishkin (2002) discuss word-internal stress patterns in Hebrew in an Optimality Theoretic (OT) framework. They argue that the regular pattern of stress in Hebrew is word-final. Other constraints on prosodic structure, due to lexical accent specifications on particular stems and suffixes, interact with the default rhythmical

pattern to result in the emergence of non-default stress patterns. They follow Graf (2000), who proposes that when there are no lexical properties specified on the stem and/or for the suffix, a default stress rule is activated, by which primary stress is assigned to the rightmost accented syllable in the noun (cf. Bat-El's (1993) formulation of similar rules in the framework of metrical theory). However, lexical stress is applied before phonological stress, and has priority over the default phonological stress, so non-final occurrences of stress in nouns thereby arise. When both the stem and its suffix are lexically specified for stress, the right-most element "wins" and final stress is attested.[67]

Graf & Ussishkin (2002) discuss stress assignment in nouns in detail, noting that all regular bisyllabic nouns display final stress, regardless of the weight of the first syllable. They also observe that Modern Hebrew has no prosodic structure below the level of the syllable,[68] and that Hebrew as a whole tends to have a CVC syllable in word-final position. Nouns with three syllables have a final primary stress, with a secondary stress on the initial syllable. Morphological complexity does not play any role for regular stress assignment. Four-syllable nouns have a final primary stress and secondary stress on

---

[67] Bolozky (2000) claims that although the majority of suffixes attract stress to themselves, there is still a minority which attest a tendency in Modern Hebrew to maintain the original position of the stress even when inflectional suffixes are appended, in contrast to the systematic stress shift to the suffix in most words. The result in a penultimate stress, as in examples (i) or even further to the left, as in (ii).

i.	modiín	modiíni
	intelligence	of-intelligence

ii.	šlúmper	šlúmperi	šlúmperiut
	slob	being a slob	'slob-ness'

Bolozky argues, following Noyman (2000), that the majority of cases follow the prescriptive normativism which has its roots in Old Hebrew, while there exists a second tendency in Modern Hebrew to opt for stable stress. There are two main reasons: first, in order to maximize the transparency of the stem by maintaining a fixed stress even in inflectional derivation. Second, the relative weight of the stem to which a suffix is attached is heavier in these cases, and this inhibits stress from shifting. Bolozky nevertheless notes that the tendency to use the non-default stress, as in the examples above, is mostly in very colloquial words or in names; therefore he suggests that the use of any deviating stress marks informality or familiarity.

[68] Graf (2000) notes that Modern Hebrew is uniformly considered a quantity insensitive language: stress assignment does not distinguish between open and closed syllables (the weight of the final syllable does not affect stress placement).

the antepenultimate syllable. (It is difficult to find examples of words longer than four syllables, even nouns (except for loan words), because of prosodic size restrictions, which cause vowel deletion; see also footnote 63.)

Furthermore, Graf & Ussishkin (2002) offer a unified analysis of the verbal and nominal systems in Hebrew. They note that in the verbal system, affixes (both prefixes and suffixes) are often lexically marked and therefore affect stress assignment. Suffixes in Hebrew are always inflectional, while prefixes may be derivational (determining a derivation class, a *binyan*) or a fused form of a derivational and an inflectional affix. They review bisyllabic, trisyllabic and quadrisyllabic forms and conclude that in all cases, both for nouns and for verbs, the main stress falls by default on the final syllable, and secondary stress is iterative, assigned to every other syllable to the left of the main stressed syllable, in accordance with Graf's (2000) analysis.[69] However, there are exceptions to the regular stress pattern, as was noted above, due to lexical specifications of stems and affixes. The verbal system manifests a systematic deviation from the default rule for some derivation classes: e.g., the *hif'il* derivational class (*binyan*) systematically has penultimate stress in the past tense forms, as in the examples in (4), where the stem is *hisgir*.

(4)	a. hisgárti frame (1, sg, msc, past tense)	cf. masgír frame (1, sg, msc, present tense)
	b. hisgíra frame (3, sg, fem, past tense)	cf. masgirá frame (3, sg, fem, present tense)

[69] Graf & Ussishkin (2002) point out that Hebrew words are templatic: stem material stays unparsed (vowel deletion occurs) in order to meet rigid prosodic size restrictions (templates). This is more often relevant in the verbal system than in the nominal system. For this reason, it is very difficult to find trisyllabic verb forms, or any word longer than four syllables (with loan words as possibly the exception).

Graf & Ussishkin show that in the verbal system, stress falls on the rightmost syllable of the stem, which is at the same time the right edge of the prosodic word. When a stem is followed by a consonant-initial suffix, stress is penultimate, since consonant-initial suffixes fall outside the prosodic domain. When the stem is followed by a vowel-initial suffix, stress is final, since vowel-initial suffixes must be part of the prosodic word in order to be parsed. In all other cases, the default rhythmical rule assigns main stress to the final syllable, as argued in Graf (2000).[70]

Becker (2002, 2003) investigates the phonetic realization of stress in Hebrew. He notes that in Hebrew, there is no underlying tone, and there is no phonemic vowel-length or consonant-length distinction (see also Graf & Ussishkin (2002) for a similar claim about the lack of distinction between long and short vowels). These two aspects of pronunciation can therefore be used as correlates of stress in his analysis. Becker shows, based on experimental and corpus data of unplanned, every-day-like speech, that stress in Hebrew is realized by both vowel length and a high tone.[71] He notes that vowel length straightforwardly marks the stressed syllable: based on phonetic measurements, he shows

---

[70] It has been noted in Becker (2002, 2003), that the analysis in Graf & Ussishkin (2002) is based on the occurrence of words in isolation. Becker argues that when words are examined at the phrasal level, stress is assigned not to the prosodic word but to the lexical word. He compares example (i) in isolation (a) with its occurrence before words with different stress patterns: with final stress (b) and with penultimate stress (c).

(i) a. roní:t
Ronit

b. ronit halxá
Ronit went

c. roní:t ká:ma
Ronit rose

In (a) the isolated stress location for the word is attested. However, while in (c) stress on *roni:t* remains, in (b) there is a shift in stress, leaving *ronit* stressless. Becker argues that it is not the adjacency of the two high tones, but the adjacency of the two stressed syllables that blocks stress shift (c), while when *roni:t* is followed by an unstressed syllable, stress shifts (b). He argues that the right edge of the foot of a head should therefore be aligned with the right edge of the lexical word.

[71] Becker's data on Hebrew tonology comes from two sources: (1) a series of laboratory experiments with four native speakers; (2) recordings of radio talk shows such that the speakers were judged to be native speakers, and the speech was fluent, unplanned and every-day-like.

that regardless of the syllable structure of the word, the stressed vowel in the word is twice the length of the vowels of unstressed syllables in that word. This is the case, for example, in *yé:·led* 'child' and in *ba·ló:n* 'baloon'; the two words have different syllable structures, but in both words the stressed syllable is twice the length of the stressless syllable. He further shows that a stressed syllable is marked by a high tone. Specifically, a high tone appears in a stressed syllable when it is final or penultimate in the phrase. When the stress is farther to the left in the phrase, the high tone appears one syllable after the stress. Becker claims that tone shift is a productive phenomenon in Hebrew: in principle, Hebrew assigns high tone to every stressed syllable in a phrase, and a rule pushes the high tone one syllable to the right in some contexts.[72]

3.4 Summary

The aim of this chapter was to draw together the findings available to date on Hebrew prosody. Among the few sources for this review, one major work on intonation (Laufer, 1987) provided detailed information on 6 nuclear tones in Hebrew, and the most common intonation contours in the language. In addition, several works have investigated acoustic cues for prosodic boundaries in Hebrew (in both planned and spontaneous speech), which include a change in pitch, a lengthening of the final syllable and a physical pause. Finally, the chapter reviewed stress assignment within the Hebrew nominal system, which is pertinent to the CS nominals in the present study.

[72] Becker (2002) lists several factors which block tone shift. First, when stress is on the final syllable of the word, or when the syllable that follows the stress is final in the word, tone shift is blocked. Second, in the presence of another stressed syllable, high tone is blocked from shifting to the right. Lastly, shifting is optional when the shift crosses into a following word.

The information in this chapter will serve as the basis for a ToBI-like notation system (Chapter 5, §5.2.5.1), which will be used to analyze elicited utterances in the production experiment of this dissertation (see Chapter 5). The advancement of our understanding of Hebrew prosody is encouraging; this dissertation makes some contributions to this area of study.

CHAPTER 4

MATERIALS AND GENERAL DESIGN OF EXPERIMENTS

This dissertation examines the relationship between prosodic phrasing and the interpretation of an ambiguous RC attaching to a preceding complex nominal, as presented in the examples in Chapter 1 above and repeated in (1).

(1) (ha-)me'amen (šel) ha-mit'agref še-paraš (le'axar ha-taxarut)
(the-)coach (of) the-wrestler who-retired (after the-fight)
'The coach of the wrestler who retired (after the fight)'

Three experiments were designed to investigate the RC construction in Hebrew: A production experiment established the prosodic phrasing of sentences containing the RC construction when disambiguated contextually; a listening experiment provided data on the perception of RC sentences uttered with varying prosodic phrasing; a silent reading experiment revealed attachment preferences of the ambiguous RC construction in Hebrew. Throughout the experimental component of this dissertation, two main factors are manipulated: Complex nominal type and RC length. While the three experiments differ in other aspects of their design, as will be discussed separately for each experiment in the following chapters, they all share these two general design factors.

The first factor is complex nominal type: The complex nominal preceding the RC is either an FS or a CS. As has been discussed above (see Chapter 2), these two closely related complex nominals have quite different syntactic and phonological properties, despite sharing a common meaning. The unique phonological properties of the CS

suggest that the prosodic phrasing of sentences containing the CS might differ from those with the FS. It is widely agreed in the literature that the CS is a single phonological word. In contrast to the CS, each of the nouns in the FS constitutes a separate phonological word. In consequence, different prosodic phrasing patterns are possible for each of the constructions under investigation, with potential effects on attachment interpretation.

The second factor is the length of the RC. RC length effects on attachment interpretation have been observed in numerous other languages (Standard (Egyptian) Arabic: Abdelghany, in preparation; Croatian: Lovrić, 2003; Dutch: Wijnen, 2001; English/Spanish: Fernández, 2000/2003; Bradley et al., 2003; French: Pynte, 1996; Pynte & Colonna, 2000; German: Hemforth & Konieczny, 2002; Augurzky, 2005; Japanese: Kamide et al., 1998; Brazilian Portuguese: Lourenço-Gomes, 2002; Maia et al., 2004; among others). Two RC lengths are tested in Hebrew: a 'short' RC (containing one minor phrase) and a 'long' RC (containing three minor phrases). For each item, the long RC was an extended version of the short RC, starting with the same minor phrase but continuing with two additional minor phrases; care was taken that the additional material did not substantially alter the meaning of the RC. For the sentence in (1), for example, the long RC *še-paraš le'axar ha-taxarut* 'who retired after the fight' is an extended version of the short RC *še-paraš* 'who retired', containing 'boring' material such as time and place adjuncts and thus adding little additional information in this sentence, given its coaching/wrestling scenario. The extent to which prosodic phrasing interacts with RCs of various lengths may prove significant in evaluating the attachment biases for the two types of complex nominals. In addition, to find significant effects of RC length would validate the sensitivity of the experimental protocol.

4.1 Materials

Twenty-four items were constructed to investigate the RC construction in the set of experiments reported in this dissertation. The target sentences contained one or other of the two types of complex nominals, followed by an RC. The full list of experimental items is presented in Appendix A-1. All target sentences had the complex nominal in direct object position, with exactly two minor prosodic phrases preceding it: a definite subject consisting of a single noun (with an affixed determiner where appropriate) and then a verb in past tense. It was important that all target sentences have an identical syntactic structure for the material preceding the target construction so that any variations in the prosodic structure may be attributed solely to variations in the factors investigated. The subject-verb sequence also ensured that the materials preceding the target construction are not unusually long so they would not be likely to induce a prosodic break before the complex nominals.[73] A schematic representation in English is presented in (2).

(2) (The-)N0 V_{past} N1 (of) N2 RC

Null subjects were avoided in order to maintain the requirement for two minor phrases preceding the complex nominal containing the RC, as were pronouns and low-salience generic terms such as *the man*, *people*, etc. The accusative verb was consistently followed by an accusative Case marker *'et*, a monosyllabic and unstressed grammatical morpheme which cliticizes onto the object, the prosodic word that follows it in the

[73] The sequence of an overt subject followed by a verb was the minimum structure required in order to create a grammatical and varied battery of experimental sentences. A covert subject is possible in Hebrew but only under restricted conditions and therefore would have created an unnatural battery of sentences.

sentence, and therefore does not carry any prosodic weight.[74] The gap in the relative clause was always in subject position, i.e., the relative pronoun in all experimental sentences referred to the subject of the relative clause. It was important to use the same type of RC for reasons of consistency; the subject gap RC was chosen because it requires the least processing load,[75] and also because it is the form used in previous work on other languages, making it possible to compare these experiments' data to existing data.

In Hebrew, the verb within the RC agrees with its subject, which in these materials is the trace, which agrees with the head noun that the RC modifies in gender and number. Therefore, the two nouns within the complex nominal were matched for gender and for number, such that both were grammatically licit candidates for the RC attachment. Both were in the singular form,[76] and were either both masculine or both feminine. The two nouns were also matched for animacy and definiteness, to minimize any semantic biases.[77] One half of the experimental materials contained animate complex nominals and the other half contained inanimate complex nominals; using both animate

---

[74] Prepositional verbs (e.g., *diber im* 'talked with') were avoided in target items because of the possibility that they might affect prosodic boundaries between the verb and the following nominal. Such verbs were used in some of the filler items.

[75] Several studies have shown that object relative clauses are associated with more processing difficulty than subject relative clauses, e.g. for English: Ford, 1983; King & Just, 1991; King & Kutas, 1995; for German: Mecklinger, Schriefers, Steinhouer & Friederici, 1995; Schriefers, Friederici & Kühn, 1995; for Dutch: Frazier, 1987; Mak,Vonk & Schriefers, 2006.

[76] Plural nouns were avoided because it was found, for English, that results differ in attachment preferences for singular and plural nouns (Fernández, 2000/2003): In both English and Spanish, there was a general shift towards high attachment when both attachment sites were plural nouns. Fernández found that the preference to attach low in English in her examples with singular nouns disappears when the attachment hosts are plural. She notes, "Sentences with singular complex NPs induce outcomes closer to the standardly reported picture of low attachment preferences for English […], while those with plural complex NPs depart more drastically from it" (p. 267). It is uncertain whether Hebrew exhibits sensitivity to number features in the complex nominal in a similar way, but in order to obtain data that can be compared with the standard findings for English and other languages, it was decided to use the singular form of the complex nominal.

[77] Desmet et al. (2002, 2005) have shown that in a Dutch corpus study and on-line reading task, animacy and concreteness attract RC attachment. Thus, in cases such as *the decisions/documents of the president*, the animate noun (*president*) attracts the RC more strongly than abstract (*decisions*) or concrete (*documents*) inanimate nouns.

and inanimate complex nominals ensured that the results were not biased by animacy considerations and could be generalized. The complex nominal as a whole was definite, as in prior studies on other languages, which have used only definite nominals. In constructing the target items for these experiments, an attempt was made to select nouns that were approximately equally frequent;[78] this ensured that the two nouns were equally likely candidates for modification by the RC, in terms of their salience.[79]

Different types of relations were expressed by the complex nominal, including kinship relations and possessive relations. Complex nominals that contained both an abstract and a concrete noun were excluded (all inanimate complex nominals were concrete, see footnote 77), as were deverbal or process nouns with an associated argument structure (e.g., *the destruction of the city*, *the review of the poem*).[80] Part-whole or inherent possessive relations within the complex nominal were permitted but only when the meanings of the two nouns were sufficiently separate that each could clearly be distinguished as a candidate for the modification of the RC (e.g., not *the branch of the tree that bloomed*, in which N1-attachment interpretation entails N2-attachment

[78] At the time when the materials were constructed, no list of word frequencies existed in Hebrew, thus the author of this dissertation made an intuitive judgment regarding the frequency of the nouns used. Two Hebrew databases for word frequency were compiled since: "The word-frequency database for printed Hebrew" by Ram Frost and Dave Plaut (http://word-freq.mscc.huji.ac.il/index.html), and "The Hebrew Word Frequency Database" by Alexandra McCauley (http://homepages.inf.ed.ac.uk/v1amccau/). The lack of use of diacritics (*nikud*) in these compilations makes many words ambiguous between several meanings, thus making it difficult to determine the actual frequency of one word meaning. Also, the tendency of most prepositions in Hebrew to cliticize onto the following noun creates many entries for the same noun, and it is not clear whether frequency count should include all or only some of these occurrences.

[79] Pynte & Colonna (2000) found that in French, the relative frequency of the two nouns influences attachment choice: an RC tends to adjoin to the low-frequency noun in cases of unequal frequencies between the two nouns.

[80] Assuming that deverbal/process nouns share the argument structure of the correlated verb (Argaman & Pearlmutter, 2002), in examples like *the destruction of the city*, N1 is considered a deverbal noun and N2 is an argument of the deverbal noun, and as such is within the semantic domain of N1. Gilboy et al. (1995) have shown that when N2 serves as an argument of N1, it is less likely to attract RC attachment.

interpretation and vice versa). Other than the RC itself, there was no modification of the nouns within the complex nominal.

Care was taken to ensure that there was no linkage of meaning between the matrix verb and the relative clause meaning, in order to avoid biasing the RC-attachment (e.g., in the sentence *The doctor cured the son of the colonel who died*, the high attachment reading is unavailable since the matrix verb *cured* implies that the son is unlikely to be dead). No commas were used between the complex nominal and the RC in order to avoid systematic pausing induced by the comma, which could make length manipulations less sensitive and affect attachment choice.[81] The sentences were fully acceptable without commas.

Two questionnaires pre-tested the materials in order to minimize any semantic biases that might affect the results of the three main experiments. It was not essential that the materials for the main experiments be totally unbiased between the two potential interpretations, since the important results concerned effects of the factors RC length and nominal form on items that were matched for content. The purpose of the pre-tests was to exclude any extreme biases, which might otherwise have obscured subtle influences of those factors of interest. The two pre-tests are described in details below; the first pre-test (presented in §4.2) was a rephrasability assessment designed to ensure that the complex nominal without RC was equally natural, acceptable and plausible when used in both the CS and FS forms; the second assessment test (described in §4.3) confirmed that the RC is

[81] Steinhauer & Friederici (2001) found that for German RCs which are preceded by a comma, the RCs are predominantly interpreted as attaching high. They propose that a comma is a marker of a prosodic boundary, and as such it biases the interpretation. The addition of a comma in Hebrew could also potentially make the RC a non-restrictive relative clause, which was not the intended interpretation of these sentence materials.

equally plausible as a modifier of either noun of the complex nominal in both the CS and FS, and no semantic considerations bias the RC attachment towards one of the two hosts.

4.2 Rephrasability assessment

The purpose of the rephrasability assessment was to establish that a given item is acceptably used in both the FS form and the CS form in Hebrew, and that Hebrew speakers do not tend to prefer the use of one form over the other. The questionnaire tested 40 complex nominals, of which 24 optimal candidates were chosen for use in the main experiments of the dissertation. The items were presented within declarative sentences identical to the linear structure in (2) above, but without the RC which creates the ambiguity. An example of a target item is presented in (3), with an FS version (3a) and a CS version (3b).

(3) a. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref.
the-fans admired ACC the-coach of the-wrestler

b. ha-'ohadim he'ericu 'et me'amen ha-mit'agref.
the-fans admired ACC coach the-wrestler

These items were mixed with 40 filler items, for a total of 80 sentences. Filler items were complex nominals which were intuitively judged as favored either in the FS form or in the CS form. For FS favored, we used complex nominals that are clearly unnatural when used in the CS form.[82] For CS favored, we used frozen CS forms

[82] The FS favored filler items were judged by the experimenter as intuitively implausible and cumbersome, although felicitous, when used in the CS form. The sentence in (i) is an example of a CS form for an FS favored item.
(i) ha-merkazanit dibra im psixolog ha-pacyent
the-receptionist talked with psychologist the-patient
'The receptionist talked with the psychologist of the patient.'

(compounds) which lose their fixed meaning once rephrased into the FS form and consequently result in semantically incoherent FS forms; therefore, they can only be interpreted as 100% CS favored (see discussion of the CS forms vs. compounds in Chapter 2, §2.4). The CS fillers were used not only to camouflage the target items, but also as an accuracy measure for recognizing participants who are not performing well on the task. The complete list of experimental and filler items is presented in Appendix A to Chapter 4, with experimental sentences chosen for the main experiment marked by an asterisk (*).

4.2.1 Design

The questionnaire included a total of 80 sentences: 40 experimental items and 40 filler items. A single item included two versions: one in CS form and another in FS form, which were presented one below the other. Half of the experimental items were animate complex NPs and the other half were inanimate complex NPs. For filler items, half were complex nominals favoring CS and the other half were complex nominals favoring FS, as described above. The 80 sentences were divided into 5 blocks; each block included 16 sentences: 4 experimental items which were animate, 4 experimental items which were inanimate, 4 filler items favoring CS and 4 filler items favoring FS. Pseudo-randomization ensured that there were no more than two consecutive sentences of the same type (out of the 4 types of sentences), no more than 3 consecutive experimental items (animate and inanimate items) and no more than 3 consecutive filler items (CS- and

---

The results supported this intuition. There was an overall preference of 71.7% for the FS form for this set of filler items. As discussed in Chapter 2, §2.1, there is no explanation in the literature, to my knowledge, as to why the CS forms of certain items seem implausible to native Hebrew speakers, despite the fact that a different pair of words closely related to the implausible pair and conveying the same relation is judged as plausible.

FS-favored). The questionnaire was divided into two parts, with a break in the middle at which participants were encouraged to pause. There were 2 filler items at the beginning of each part (a sequence of CS favored - FS favored), and at the end of each part. The rest of the sentences for each block were in pseudo random order.

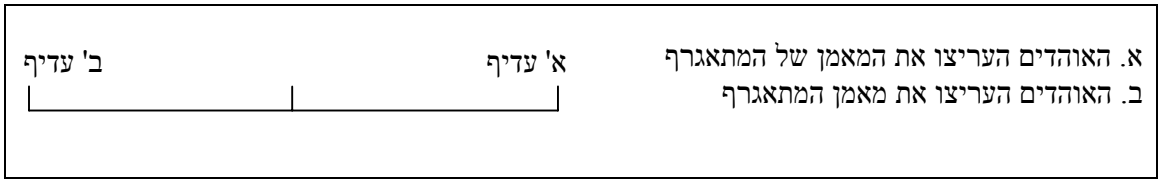
The design manipulated two factors: The first was order of presentation: Participants either saw the FS first and CS second, or the reverse order. This ensured that the mode of presentation of the two complex nominals (whether the pair is presented with FS before CS or vice versa) has no effect on the interpretation. The second design factor was order of encounter: the 5 blocks of the pretest were ordered such that the first two blocks in one questionnaire were the last two in another questionnaire; the last two blocks in one list were the first two blocks in other lists. This rules out order effects, e.g., that participants are more attentive during the first half of the questionnaire compared with a more sloppy performance in the second half of the questionnaire. An effort was made to avoid obvious discourse connections between items in adjacent or nearby positions in the sequence, so that participants could make their judgment on each item on its own merit and not link it with a recently encountered item.

4.2.2 Procedure

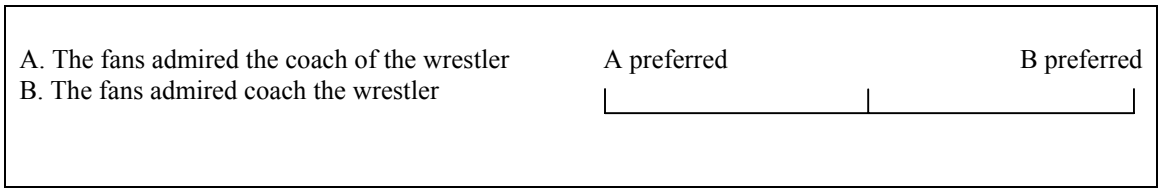
The questionnaire was presented as a booklet. Each page contained one target item, presented in two versions, one containing a CS and the other an FS. The two alternatives were presented one below the other, and preceded by the equivalent of “A” and “B” in Hebrew. Next to the two sentences, there was a scale, and participants were asked to mark on the scale whether there is a difference in the *svirut* (plausibility) and

kvilut (acceptability) of the sentences. These terms were used more or less interchangeably, with the intention of allowing the participants some freedom in interpretation with the main emphasis on whether it would occur in normal usage of the language. They were asked to assess whether sentence A was more plausible and acceptable than sentence B or vice versa, or whether they were equally good (*tovim be'ota mida*) according to their intuitions. An example in Hebrew for the sentence in (3) is presented in (4a) and the English translation in (4b).[83]

(4) a.



b.



The scale was 100 millimeters (mm) long, thus facilitating the measurement of the mark made by participants (the length of the scale in (4) is reduced from is original size for presentational purposes). Participants' marks were calculated as a measurement starting at 0 mm, which corresponded to the CS favored edge of the scale, and ending at 100 mm, which corresponded to the FS favored edge of the scale; thus, the middle point in the scale, which corresponded to judgment of equal acceptability for both FS and CS,

[83] Due to the right-to-left orthography in Hebrew, the Hebrew example in (4a) is a mirror image of the English example in (4b), with the sentences presented on the right and the scale on the left of the page.

was at 50 millimeters. This measuring strategy enabled the use of both sides of the scale as a starting point for measuring, regardless to whether CS favored was on the right or left edge of the scale (which varied in the different experimental lists). The task of the experiment was not disguised; participants were specifically asked to mark their preference. However, they did not know that there were different sentence-types and that the experimenter was interested in only a few of them.

The instructions to the questionnaire included two examples and 4 practice sentences. The two examples in the instructions included a CS favored (compound), and a more moderate FS favored item, so that participants saw visually that they could mark their judgment in any place along the scale. The practice items were diverse, and included 2 filler items and 2 items in the form of the experimental materials. The two experimental items used in the practice session followed each other in order to help participants realize that it is possible to mark judgments at the same place for two consecutive items. The items used in the practice session were not used in the final analysis of the pretest. The booklet contained 3 filler items used as covert practice items prior to the 80 sentences described above. The instructions and practice items for the rephrasability questionnaire are presented in Appendix A to Chapter 4.

4.2.3 Participants

Sixteen native speakers of Hebrew completed the pretest successfully. They were employed at the time at the Consulate General of Israel in New York, and were all in their 20s and early 30s. All participants had been in the US for a short period of time (not more than 2-3 years) and despite their presence in an English speaking country, they all

testified that Hebrew was the main language they used both at work and in their private lives. Thus although they were drawn from a different pool, their characteristics matched the population targeted in the three main experiments reported in this dissertation.

4.2.4 Item selection procedure

Twenty-four target items out of the 40 presented in the acceptability/plausibility questionnaire were used in the main experiments. 12 items were chosen out of the 20 animate items, and 12 were chosen of the 20 inanimate sentences, in order to keep a balance between animate and inanimate sentences. The 24 chosen sentences were judged as best candidates in terms of equal plausibility and acceptability in their respective FS and CS forms. This eliminated any sentence with an average judgment of less than 25% and higher than 75%; thus the accepted sentences were those that were judged as closest to the midpoint (50%) between FS favored and CS favored.

Overall, the judgment scores for the chosen items ranged from 29.4% to 68.4%, with the low end characterizing an item which has a preference for CS use and the high end characterizing an item with preference for FS use. For animate items, the range of chosen sentences was from 29.4% to 68.4%, with a mean of 52.9%; for inanimate items, the range was from 31.3% to 57.8%, with a mean of 41.7%. Thus the rephrasability assessment ascertained that the items used for the experiments reported in the following chapters are judged as plausible and acceptable for use in both the FS and CS forms, and that the FS and CS forms of these items can be rephrased into one another.

4.3 RC attachment assessment

The RC attachment assessment was designed to establish that both nouns comprising the complex nominal are equally plausible as hosts for RC attachment. It tested whether the ambiguous sentences used in the three main experiments are grammatically acceptable and semantically plausible in both of their interpretations, thus avoiding any bias of the RC towards a particular noun within the complex nominal on plausibility grounds. The primary reason for balancing semantic/pragmatic plausibility was to permit a clear view of the interpretation of these constructions when meaning is not a relevant force.

4.3.1 Design

The questionnaire assessed the 24 experimental items chosen for the three experiments, which included 12 animate and 12 inanimate sentences, as described above. Each experimental item was tested in both the CS and FS form, in order to rule out the possibility of an intrinsic bias towards RC attachment to one of the nouns in one form but not in the other. If a semantic bias were found in the materials, it would show up in both FS and CS variants of a sentence; if there were an intrinsic attachment bias in one of the forms, it would show up in one variant but not in the other.

In order to avoid repetition of the same item in both FS and CS forms within the same questionnaire, two lists were prepared: in one list half the experimental sentences were presented in the FS form and the other half in the CS form; in the second list the reverse form was presented for each experimental item. Both RC length variants were tested, but the design factor was not fully crossed: half the experimental sentences were

tested with a short RC while the other half were tested with a long RC. The RC length for each item remained the same in the two lists of the questionnaire to allow for a direct comparison. There was no reason to expect RC length to affect the assessment of these sentences, since the materials were constructed such that the added material in the long RC does not add crucial information to the sentence, as described above (§4.1).

The 24 experimental items were mixed with 24 filler items. The fillers were included in order to induce a varied range of responses. As in the experimental items, fillers had a 1:1 ratio of short to long RCs. The filler items had an identical sentence structure to that of experimental items, with an RC attached to a complex nominal, but attachment interpretation was not expected to be equally plausible. The filler sentences were divided into two types. 12 fillers were compounds which are either opaque, or transparent but lexicalized (see Chapter 2, §2.4). For these, N2 interpretation was implausible, while N1 modification could only be plausible as modification of the compound as a whole. These were presented in the CS form in both lists (once used with FS form, they lose their lexicalized meaning and often become unintelligible). For the 12 remaining fillers, RC was semantically and pragmatically biased towards one noun within the complex nominal (in half of these, N1 interpretation was biased and in the other half, there was a bias for N2 interpretation). These appeared in both FS and CS between the two lists, similar to experimental items.

The 48 sentences (24 experimental and 24 filler items) were presented in a pseudo-randomized order which divided them into 4 design blocks, each including 3 animate target items, 3 inanimate target items, 3 compound filler items and 3 semantically/pragmatically biased filler items. FS and CS items were intermingled within

each list. The pseudo-randomization ensured that there were no more than 2 consecutive experimental items, and that each list began and ended with a filler item. The rest of the sentences were in random order. The order of sentences was identical for the two lists, which only varied in the form of a complex nominal (i.e., in one list an item was presented in FS form and in the other the same item was presented in CS form, and vice versa). The list of experimental items used in the RC attachment assessment, which were also the items used in the three experiments of this dissertation, is presented in Appendix A-1 of the dissertation; the filler items used only in the RC attachment assessment test are shown in Appendix B to the present chapter.

4.3.4 Procedure

The assessment test was presented as a paper and pencil questionnaire. Each item was presented ambiguously, followed by two brief sentences indicating alternative interpretations of it. Next to each interpretation was a 5-point scale and the task required participants to mark on the scale whether the interpretation is syntactically and semantically acceptable/plausible (5) or unacceptable/implausible (1). An example of the task (for example (1) above, in its FS variant) is presented in (5a) in Hebrew, and in the English translation in (5b). For experimental items, it was expected that both alternative interpretations would be judged grammatical/plausible, while filler items were expected to have varied responses, as discussed above.[84]

[84] As in the examples in (4), the Hebrew version follows the right-to-left orthographic rules of Hebrew, therefore the sentences are displayed on the right and the scale on the left (a mirror image of the English version).

(5) a.

האוהדים העריצו את המאמן של המתאגרף שפרש.					
1	2	3	4	5	המאמן פרש.
1	2	3	4	5	המתאגרף פרש.

b.

The fans admired the coach of the wrestler who retired.					
The coach retired	1	2	3	4	5
The wrestler retired	1	2	3	4	5

Three expert judges were chosen to participate in this test (the writer of this dissertation was not one of them). They were asked to assess the experimental sentences in order to determine the legitimacy of the two interpretations of the ambiguous sentences. All judges were native speakers of Hebrew and have a background in linguistics. The same 3 judges assessed both lists, which were presented in two different meetings separated by 3 weeks' time.

The instructions to the assessment questionnaire clearly described the ambiguity of the RC construction, and the task of assessing whether the two alternative interpretations are acceptable and plausible syntactically and semantically. The instructions emphasized that the judges should use their sensitivity to language in order to make a considered judgment for each interpretation, not just rely on a quick first impression (because the judges were familiar with psycholinguistic experiments which emphasize the need for a quick intuitive response, it was important to note that their responses in this assessment task needed to be carefully considered.) The instructions

also noted that judgments could be marked on any point on the scale, and that the judgments for the two interpretations of a sentence were not required to be the same; it was expected that while for some sentences both alternatives would be acceptable, for other sentences one (or both) of the interpretations would be unfeasible for different reasons. The instructions included one experiment-like example, in which both alternative interpretations are plausible. The instructions were followed by a practice session, which included 4 items, before the list of sentences was presented. Two practice items were experiment-like sentences (one with a short RC and the other with a long RC); one had a compound and the fourth was a filler sentence in which one of the attachment interpretations was semantically/pragmatically implausible. The instructions and practice items are presented in Appendix B to this chapter.

4.3.3 Results

The data for the experimental items show that N1 and N2 attachment was judged highly plausible, and roughly equally so, both in the FS and in the CS. Overall, there was no clear preference for one attachment host over the other, either in the FS or the CS, and both complex nominals had high rates of judged plausibility when N1 and N2 were hosts. The agreement rate among the three judges was high; in 88% of the cases there was no difference or a difference of only 1 scale-point in their markings. Table 1 shows the results of the RC attachment assessment in terms of likelihood of N1 and N2 as attachment hosts.

Table 4-1: Percentage of plausibility/acceptability of N1 and N2 in CS and in FS.

	N1	N2
FS	86.8	96.2
CS	93.4	96.2

When comparing CS and FS, the plausibility of N1 attachment was fairly similar in both complex nominals (86.8% and 93.4% respectively), and the plausibility of N2 attachment was identical in both constructions (96.2%). An item-by-item analysis revealed a few cases in which there was a slight preference for either N1 attachment or N2 attachment, but these were the exception and removing them from the grand mean calculation did not change the results in any crucial way.[85]

Semantically biased fillers showed more varied responses, as expected; responses were similar whether the filler was presented with a CS or an FS. The bias in attachment plausibility was in the direction predicted, i.e., for sentences designed to have an N1 attachment bias, the plausibility of N1 attachment was 95.8% and 100% for FS and CS, respectively, while N2 was clearly dispreferred (the plausibility of N2 attachment was 55.6% and 48.6% for FS and CS, respectively). For sentences in which attachment was biased towards N2, the plausibility of N2 attachment was very high (98.6% for FS and 100% for CS) while N1 attachment was clearly less plausible (43.1% for FS and 56.9% for CS).

[85] Three items had a slight difference in N1 and N2 attachment plausibility (in two there was a small difference favoring N2 attachment and in one there was a small preference for N1 attachment). Removing these three mild exceptions from the grand means calculation showed that the FS had N1 and N2 as likely attachment hosts 88.9% and 97.2%, respectively, and CS had N1 and N2 as plausible attachment hosts 93.6% and 96.4%, respectively. Thus, they do not vary in any important way from the grand means reported above for all items combined. It is interesting to note that these three items were all inanimate sentences with the long RC variant.

Compound fillers showed categorical rejection of N2 as a plausible host, as predicted, with 0 plausibility for the N2 attachment interpretation. This was expected, since N2 attachment in most compound items was ungrammatical due to gender/number mismatch (as is often the case in compounds). The 3 cases in which N2 attachment was grammatical also exhibited categorical rejection of N2 interpretation; it is likely that N2 interpretation is rejected not merely due to semantic implausibility but due to the fact that N2 modification (by adjectives, numerals or relative clauses) is syntactically blocked in compounds (see Chapter 2, § 2.4).[86] N1 attachment, on the other hand, was fully accepted in all compound items. This was surprising, since the interpretation option presented to participants included only N1 of the compound, not the whole compound, and therefore the N1 interpretation was predicted to be rejected as well since the meanings can be quite different (e.g., *ha-melcar hegiš 'et tapuxey ha-'adama še-bušlu be-rotev meyuxad* 'the waiter served apples the-ground (potatoes) that were cooked in a special sauce'). It is possible that in some of the cases, N1 of what was taken to be a lexicalized compound was still transparent enough for modification by itself, unrelated to the interpretation in the context of the compound construction (e.g., *Bracha hitya'aca im orex ha-din she-humlac be-reshimot ha-lishka* 'Bracha consulted with editor the law (the lawyer) who was recommended in the office's lists'). Alternatively, in the absence of one plausible host for an otherwise grammatical sentence, judges may have assumed that the head of the compound was being used in this questionnaire as a shorthand for the compound as a whole, rendering that interpretation plausible. In short: judges clearly made different

[86] It is interesting to note that those compound cases in which N2 interpretation was rejected on semantic implausibility grounds were judged as completely implausible, while fillers which had CS and FS complex nominals but were semantically biased towards one interpretation were judged less categorically (around 50% plausible, in contrast to the highly plausible alternative interpretation).

acceptability judgments on filler items, compared with experimental items, with more varied responses to the semantically biased sentences, in the direction predicted, and a categorical rejection of N2 attachment and acceptance of N1 attachment in the case of compounds. For experimental sentences, the RC attachment assessment confirmed that both N1 and N2 are equally plausible attachment hosts, and equally so for FS and for CS sentences.

The results of the two assessment tests provide sufficient evidence that the materials constructed for use in the three main experiments are not semantically biased towards one or the other attachment host of the complex nominal. They also ensure that both CS and FS forms are equally plausible for use in the sentences presented in this set of experiments. In addition, because the main experiments tested each of the 24 items in both the FS and CS, any residual imbalance in their plausibility could not invalidate evidence of the effects of the nominal type (FS or CS) and/or the RC length (short or long). The RC attachment assessment and the complex nominal assessment reported here provide reassurance that the materials do not contain any extreme bias, which might impede subtle influences of those factors.

Appendix A to Chapter 4: Rephrasability assessment

A-1: Materials

Below are the 40 experimental sentences used in the pretest, with the sentences chosen for the three main experiments marked by an asterisk (*). Sentences 1-20 are animate experimental sentences; sentences 21-40 are inanimate experimental sentences. These are followed by 40 filler items: sentences 41-60 are filler items with a CS favored used; Sentences 61-80 are filler items with an FS favored use.

1. ha-katevet ra'ata 'et yoec ha-nasi'
the-reporter saw ACC advisor the-president
'The reporter saw the advisor of the president.'

2. * ha-'itona'y ri'ayen 'et bat ha-šofetet
the-journalist interviewed ACC daughter the-judge
'The journalist interviewed the daughter of the judge.'

3. * Yonatan pagaš 'et ben ha-rofe'
Jonathan met ACC son the-doctor
'Jonathan met the son of the doctor.'

4. * ha-mištara kivna' 'et nehag ha-mefaked
the-police directed ACC driver the-commander
'The police directed the driver of the commander.'

5. * ha-'oraxim birxu 'et 'axot ha-kala
the-guests congratulated ACC sister the-bride
'The guests congratulated the sister of the bride.'

6. * ha-rakezet xipsa 'et nexdat ha-cayeret
the-director looked.for ACC granddaughter the-painter
'The director looked for the granddaughter of the painter.'

7. * ha-yedi'a' hefti'a' 'et te'omat ha-zameret
the-news surprised ACC twin the-singer
'The news surprised the twin of the singer.'

8. ha-se'ara hed'iga 'et šutaf ha-kablan
the-storm worried ACC partner the-constructor
'The storm worried the partner of the constructor.'

9. ha-so'adim šibxu 'et 'ozer ha-tabax
the-diners praised ACC assistant the-cook
'The diners praised the assistant of the cook.'

10. ha-balaš xakar 'et šaxen ha-kaba'y
the-investigator investigated ACC neighbor the-firefighter
'The investigator investigated the neighbor of the firefighter.'

11. ha-mišpaxa bikra 'et dod ha-xayal
the-family visited ACC uncle the-soldier
'The family visited the uncle of the soldier.'

12. * ha-horim hecdiku 'et madrixat ha-ganenet
the-parents justified ACC instructor the-kindergarten.teacher
'The parents justified the instructor of the kindergarten teacher.'

13. * ha-moda'a hex'isa 'et 'em ha-tinoket
the-note angered ACC mother the-baby
'The note angered the mother of the baby.'

14. * ha-'ohadim he'ericu 'et me'amen ha-mit'agref
the-fans admired ACC coach the-wrestler
'The fans admired the coach of the wrestler.'

15. * ha-ti'un het'a 'et sanegor ha-ganav
the-argument misled ACC advocate the-thief
'The argument misled the advocate of the thief.'

16. ha-'askan hekim 'et hanhalat ha-kvuca
the-businessman formed ACC management the-team
'The businessman formed the management of the team.'

17. * ha-'eda ti'ara 'et me'avte'ax ha-sar
the-witness described ACC security.guard the-minister
'The witness described the security guard of the minister.'

18. ha-pkak he'et 'et sgan ha-mankal
the-traffic.jam slowed.down ACC deputy the-chairman
'The traffic jam slowed down the deputy of the chairman.'

19. ha-dox mecayen 'et xaver ha-talmid
the-report mentions ACC friend the-student
'The report mentions the friend of the student.'

20. * ha-morim hergizu 'et mazkirat ha-menahelet
the-teachers annoyed ACC secretary the-principal
'The teachers annoyed the secretary of the principal.'

21. ha-'oman šixzer 'et tikrat ha-knesiya
the-artist renovated ACC ceiling the-church
‘The artist renovated the ceiling of the church.’

22. * Ruti šatfa 'et xalon ha-rexev
Ruti washed ACC window the-vehicle
‘Ruti washed the window of the vehicle.’

23. ha-ganan heška 'et ginat ha-sifriya
the-gardener watered ACC garden the-library
‘The guard watered the garden of the library.’

24. * ha-musaxnik garar 'et mexonit ha-xevra
the-mechanic toed ACC car the-company
‘The director toed the car of the company.’

25. * širi kanta 'et sefer ha-kurs
Širi bought ACC book the-course
‘Širi bought the book of the course.’

26. ha-milioner taram 'et cilum ha-pesel
the-billionaire contributed ACC photo the-statue
‘The billionaire contributed the photo of the statue.’

27. * ha-yeled šavar 'et beycat ha-tarnegolet
the-boy broke ACC egg the-hen
‘The boy broke the egg of the hen.’

28. * ha-dayar na'al 'et delet ha-dira
the-tenant locked ACC door the-apartment
‘The tenant locked the door of the apartment.’

29. Roni lixlex 'et ripud ha-mošav
Roni dirtied ACC cover the-seat
‘Roni dirtied the cover of the seat.’

30. ha-taxana acra 'et šidur ha-misxak
the-station stopped ACC broadcasting the-game
‘The station stopped the broadcasting of the game.’

31. ha-soxer hexzir 'et masax ha-maxšev
the-merchant returned ACC monitor the-computer
‘The merchant returned the monitor of the computer.’

32. * ha-šef heš'il 'et mixse' ha-sir
the-chef borrowed ACC cover the-pot
‘The chef borrowed the cover of the pot.’

33. * ha-po'el 'araz 'et tmunat ha-mizraka
the-worker packed ACC picture the-fountain
‘The worker packed the picture of the fountain.’

34. * ha-šomer hestir 'et man'ul ha-ša'ar
the-guard hid ACC lock the-gate
‘The guard hid the lock of the gate.’

35. ha-moxeret hedbika 'et 'atifat ha-matana
the-vender taped ACC wrapping the-present
‘The vender taped the wrapping of the present.’

36. * ha-'ovdim 'ixsenu 'et 'arizat ha-madpeset
the-workers stored ACC package the-printer
‘The workers stored the package of the printer.’

37. * ha-texnay herim 'et xut ha-telefon
the-technician lifted ACC cable the-telephone
‘The technician lifted the cable of the telephone.’

38. * ha-nose'a hexzik 'et yadit ha-mizvada
the-traveler held ACC handle the-suitcase
‘The traveler held the handle of the suitcase.’

39. ha-menaka kircefa 'et ricpat ha-mirpeset
the-cleaner cleaned ACC floor the-balcony
‘The cleaner cleaned the floor of the balcony.’

40. * ha-adrixal 'icev 'et miznon ha-muze'on
the-architect designed ACC cafeteria the museum
‘The architect designed the cafeteria of the museum.’

41. ha-tošavim hez'iku 'et mexabey ha-'eš
the-residents alerted ACC extinguishers the-fire
‘The residents alerted the extinguishers of the fire.’ (‘the firefighters’)

42. ha-miflaga baxra 'et roš ha-'opozicya
the-party chose ACC head the-opposition
‘The party chose the head of the opposition.’

43. ha-'iriya heznixa 'et beit ha-'avot
the-municipality neglected ACC house the-fathers
‘The municipality neglected the house of the fathers.’ (‘the elderly home’)

44. ha-tayarim mil’u 'et sde' ha-te'ufa
the-turists filled ACC field the-flying
‘The turists filled the field of flying.’ (‘the airport’)

45. Dana hit'akva be-kupat ha-xolim
Dana was.detained in-booking.office the-sick
‘Dana was detained in the booking office of the sick.’ (‘the clinic’)

46. Noam zarak 'et mišxat ha-šina'im
Noam threw ACC paste the-teeth
'Noam threw the paste of the teeth.' ('the toothpaste')

47. ha-'acan 'ibed 'et na'aley ha-sport
the-sprinter lost ACC shoes the-sports
'The sprinter lost the shoes of the sports.' ('the sneakers')

48. Kulam 'ohavim 'et gan ha-xayot
everyone love ACC garden the-animals
'Everyone loves the garden of the animals.' ('the zoo')

49. ha-mefik 'iter 'et šadran ha-radio
the-producer located ACC broadcaster the-radio
'The producer located the broadcaster of the radio.' ('the radio broadcuster')

50. ha-šoter hexšir 'et pakax ha-tnu'a'
the-policeman trained ACC inspector the-traffic
'The policeman trained the inspector of the traffic.' ('the traffic inspector')

51. ha-tove'a tafas 'et galgal ha-hacala
the-drowning.guy caught ACC wheel the-rescue
'The drowning guy caught the wheel of the rescue.' ('the life buoy')

52. Rina tikna 'et miškafey ha-šemeš
Rina fixed ACC glasses the-sun
'Rina fixed the glasses of the sun.' ('the sunglasses')

53. ha-davar roken 'et teyvat ha-do'ar
the-mailman emptied ACC box the-mail
'The mailman emptied the box of the mail.' ('the mailbox')

54. ha-melcar hegiš 'et tapuxey ha-'adama
the-waiter served ACC apples the-ground
'The waiter served the apples of the ground.' ('the potatoes')

55. ha-xadranit sidra 'et kley ha-mita
the-chambermaid arranged ACC tools the-bed
'The chambermaid arranged the tools of the bed.' ('the bed sheets')

56. ha-xogegim 'axlu 'et ugat ha-šokolad
the-celebrating.people ate ACC cake the-chocolate
'The people celebrating ate the cake of the chocolate.' ('the chocolate cake')

57. ha-dayag nika 'et xanut ha-dagim
the-fisherman cleaned ACC store the-fish
'The fisherman cleaned the store of the fish.' ('the fish store')

58. ha-hanhaga hekpi'a 'et toxnit ha-megera
the-leadership froze ACC plan the-drawer
‘The leadership froze the plan of the drawer.’ (‘the drawer plan’)

59. ha-sfina 'agna be'ir ha-namal
the-ship anchored in-town the-harbor
‘The ship anchored in the town of the harbor.’ (‘the seaport’)

60. ha-yoledet huvhela 'el xadar ha-miyun
the-woman.in.labor was.rushed to room the-classification
‘The woman in labor was rushed to the room of classification.’ (‘the emergency room’)

61. ha-te'una sikrena 'et metapelet ha-pe'ota
the-accident intrigued ACC nurse the-baby
‘The accident intrigued the nurse of the baby.’

62. ha-nof hipnet 'et sabal ha-trmila'i
the-view hypnotized ACC porter the-traveler
‘The view hypnotized the porter of the traveler.’

63. ha-magazin piter 'et calam ha-dugman
the-magazine fired ACC photographer the-model
‘The magazine fired the photographer of the model.’

64. ha-tna'im hekšu 'al xoveš ha-pacu'a
the-conditions hardened on doctor the-injured
‘The conditions made it hard for the doctor of the injured.’

65. ha-profesor hedrix 'et menate'ax ha-me'ušpaz
the-professor directed ACC surgeon the-hospitlized.man
‘The professor directed the surgeon of the patient.’

66. Gil hergi'a 'et ba'al ha-mešoreret
Gil calmed ACC husband the-poet
‘Gil calmed the husband of the poet.’

67. ha-mehandes he'esik 'et xašmela'I ha-šipucnik
the-engineer employed ACC electrician the-handyman
‘The engineer employed the electrician of handyman.’

68. ha-hacaga rigša 'et me'aperet ha-koxevet
the-play excited ACC makeup.artist the-star
‘The play excited the makeup artist of the star.’

69. ha-merkazanit dibra 'im psixolog ha-pacyent
the-receptionist talked with psychologist the-patient
‘The receptionist talked with the psychologist of the patient.’

70. ha-sportay sixek 'im 'axyan ha-kadurselan
the-athlete played with nephew the-basketball.player
‘The athlete played with the nephew of the basketball player.’

71. ha-manxe' 'ire'ax 'et yaxcan ha-rakdanit
the-host hosted ACC publicist the-dancer
'The host hosted the publicist of the dancer.'

72. ha-va'ada hexlifa 'et me'are'ax ha-politika'i
the-committee replaced ACC host the-politician
'The committee replaced the host of the politician.'

73. ha-pgiša kalela 'et meturgeman ha-diplomat
the-meeting included ACC interpreter the-doplomat
'The meeting included the interpreter of the diplomat.'

74. ha-soxnut giysa' 'et 'amargan ha-saxkan
the-agency recruited ACC manager the-actor
'The agency recruited the manager of the actor.'

75. moše xizek 'et antenat ha-televizia
Moše secured ACC antenna the-television
'Moše secured the antenna of the television.'

76. ha-masagist hekir 'et 'ortoped ha-atlet
the-masseuse knew ACC orthopedist the-athlete
'The masseuse knew the orthopedist of the athlete.'

77. ha-studentim hitya'acu im metargelet ha-marca
the-students consulted with intern the-lecturer
'The students consulted with the intern of the lecturer.'

78. yossi hemlic al 'orex ha-sofer
Yossi recommended on editor the-author
'Yossi recommended the editor of the author.'

79. ha-xom he'ik al lako'ax ha-mexonay
the-heat distressed on customer the-mechanist
'The heat distressed the customer of the mechanist.'

80. ha-mis'ada hezmina 'et šali'ax ha-yarkan
the-restaurant invited ACC delivery.boy the-greengrocer
'The restaurant invited the delivery boy of the greengrocer.'

A-2: Instructions and practice items

Hebrew

שאלון סבירות/קבילות המשפט: הוראות למשתתפ/ת

בכל אחד מהסעיפים הבאים, מופיעים שני ניסוחים של אותו משפט, המיוצגים כמשפט א' ומשפט ב'. המשימה שלך היא לקבוע האם ישנו הבדל בין משפט א' למשפט ב' בנוגע לסבירות/קבילות המשפט: האם שני המשפטים טובים באותה מידה, האם משפט א' עדיף על משפט ב' וכן הלאה.

אנא סמנ/י איקס (X) על הקו בכדי לציין את החלטתך. למשל, אם משפט א' עדיף הרבה יותר ממשפט ב', סמנ/י איקס (X) בקצה הקו תחת "א' עדיף". לעומת זאת, אם משפט ב' סביר ומקובל הרבה יותר ממשפט א', סמנ/י איקס (X) בקצה הקו תחת "ב' עדיף". אם שני המשפטים טובים באותה מידה, סמנ/י איקס (X) במרכז הקו. באפשרותך לסמן איקס (X) בכל מקום על הקו.
להלן דוגמאות:

א. החזאי צפה את מזג האויר ב. החזאי צפה את המזג של האויר	א' עדיף	ב' עדיף
א. הממצאים אכזבו את גינקולוג הנערה ב. הממצאים אכזבו את הגינקולוג של הנערה	א' עדיף	ב' עדיף

- o אנא קרא/י את שני המשפטים לפני שתחליט/י, ואז במהירות הערכ/י את הסבירות והקבילות היחסיים לשני המשפטים. סמנ/י את החלטתך, ועבור/י לצמד המשפטים הבא. כל צמד משפטים שונה מקודמו.
- o אנא סמנ/י את האינטואיציה הראשונית שלך בנוגע למשפטים. אל תחשוב/י יותר מדי! התחושה הראשונית שלך היא מה שאנו מחפשים.
- o וודא/י שקיבלת החלטה לגבי כל צמד משפטים.
- o אנא המנע/י מלחזור ולחשוב בשנית על צמד משפטים שכבר החלטת לגבם.

ינתן לך תרגול קצר על מנת שתתרגל/י לסגנון השאלות. בזמן התרגול תוכל/י לשאול כל שאלה שתרצה/י בנוגע למהלך הניסוי.

אנא הפוך/הפכי את הדף...

א. הימאי הפליג עם רב החובל ב. הימאי הפליג עם הרב של החובל	א' עדיף ב' עדיף
א. הקרקס מימן את מאלף הנמר ב. הקרקס מימן את המאלף של הנמר	א' עדיף ב' עדיף
א. הפרס שימח את נין הקשיש ב. הפרס שימח את הנין של הקשיש	א' עדיף ב' עדיף
א. הצופים זיהו את כפיל הקומיקאי ב. הצופים זיהו את הכפיל של הקומיקאי	א' עדיף ב' עדיף

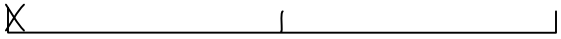
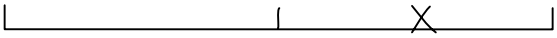
The following is a close translation of the Hebrew original; if the exact wording to the participants is important for some purpose, please consult the original.

PLAUSIBILITY/ACCEPTABILITY QUESTIONNAIRE: INSTRUCTIONS FOR SUBJECTS

In each of the following sentences, you are presented with two forms of the same sentence, identified as A and B. Your task is to indicate whether there is a difference between sentences A and B in terms of their plausibility/acceptability: are both sentences equally good, is sentence A better than sentence B and so on.

Please mark an X on the line to indicate your judgment. For example, if sentence A is much more preferred than sentence B, you would mark an X at the edge of the line under "A preferred"; conversely, if sentence B is much more plausible/acceptable than sentence A, you would mark an X at the edge of the line under "B preferred". If both sentences were equally good, you would mark an X at the center of the line. An X can be marked anywhere on the line.

Following are examples:

A. The weatherman predicted flow the air (the weather) B. The weatherman predicted the flow of the air	A preferred — B preferred 
A. The findings disappointed gynecologist the woman B. The findings disappointed the gynecologist of the woman	A preferred — B preferred 

- Please read through both sentences of the A/B pair before you judge them, and then promptly make your assessment of relative plausibility/acceptability for that pair.
- Record your judgment, and then move on to the next pair. Each pair of sentences is different from the previous one.
- Please mark your INITIAL intuition about the sentences; don't think too hard! Your first impression is what is most valuable to us. Make sure you provide an answer for each pair.
- Please avoid going back to re-work items that you have previously judged.

There will be a short practice session to familiarize you with the sort of judgment that is needed. At that time you may ask any questions you wish to, about the procedure.

Please turn the page to the other side…

A. The sailor sailed with rabbi the-roper (the captain). B. The sailor sailed with the rabbi of the roper	A preferred ⌊ ⌋ B preferred
A. The circus financed tamer the tiger B. The circus financed the tamer of the tiger	A preferred ⌊ ⌋ B preferred
A. The reward thrilled great grandchild the old man B. The reward thrilled the great grandchild of the old man	A preferred ⌊ ⌋ B preferred
A. The viewers recognized double the comedian B. The viewers recognized the double of the comedian	A preferred ⌊ ⌋ B preferred

Appendix B to Chapter 4: RC attachment assessment

B-1: Materials

Sentences 1-24 are identical to the experimental sentences presented in Appendix A-1 of the dissertation; therefore, they are not repeated here. The following are the 24 filler items used for the RC attachment assessment. Sentences 25-36 are compounds, for which either both or N2 attachment interpretation is blocked. Sentences 37-48 are semantically implausible sentences; in half of these, N1 attachment interpretation is implausible and in the other half N2 interpretation is implausible. For each category, long and short RCs are equally distributed.

25. ha-tošavim hez'iku 'et mexabey ha-'eš še-hit'akvu
the-residents called ACC distinguishers the-fire who-were.delayed
'The residents called the firefighters who were delayed.'

26. ha-miflaga baxra 'et roš ha-'opozicya še-heci'a toxnit mahapxanit
the-party chose ACC head the-opposition who-offered plan revolutionary
'The party chose the head of the opposition who offered a revolutionary plan.'

27. ha-'iriya heznixa 'et beit ha-'avot še-nisraf
the-municipality neglected ACC house the-fathers that-was.burnt
'The municipality neglected the home for the aged that was burnt.'

28. ha-tayarim mil'u 'et sde' ha-te'ufa še-nivna likrat ha-olimpiada
the-turists filled ACC field the-flying that-was.build for the-Olympic.games
'The turists filled the airport that was built for the Olympic games.'

29. Dana hit'akva be-kupat ha-xolim še-'avra le-kce ha-'ir
Dana was.detained in-booking.office the-sick that-moved to-end the-city
'Dana was detained in the clinic that moved to the end of the city.'

30. Kulam 'ohavim 'et gan ha-xayot še-nisgar
everyone love ACC garden the-animals that-was.closed
'Everyone loves the zoo that was closed.'

31. ha-melcar hegiš 'et tapuxey ha-'adama še-bušlu be-rotev meyuxad
the-waiter served ACC apples the-ground that-were.cooked in.sauce special
'The waiter served the potatoes that were cooked in a special sauce.'

32. ha-xadranit sidra 'et kley ha-mita še-huxtemu
the-chambermaid arranged ACC tools the-bed that-were.stained
'The chambermaid arranged the bedclothes that were stained.'

33. ha-yoledet huvhela 'el xadar ha-miyun še-nimca be-merxak
the-woman.in.labor was.rushed to room the-classification that-is.located in-distance

rav mi-beyta
great from-her.house

'The woman in labor was rushed to emergency room that is located a great distance from her house.'

34. ha-manxe ire'ax 'et roš ha-'ir še-hitpater
the-host hosted ACC head the-town who-resigned
'The host hosted the mayor who resigned.'

35. ha-'iriya hilela 'et beyt ha-sefer še-nexnax
the-municipality praised ACC house the-book which-was.christened
'The municipality praised the school which was christened.'

36. Bracha hitya'aca im orex ha-din she-humlac be-reshimot ha-lishka
Bracha consulted with editor the-law that-was.recommended in-lists the-office
'Bracha consulted with that lawyer who was recommended in the office's lists.'

37. ha-katevet ra'ata 'et yoec ha-nasi' še-nivxar le-kehuna nosefet
the-reporter saw ACC advisor the-president that-was.chosed for-term additional
'The reporter saw the advisor of the president who was chosen for an additional term.'

38. ha-so'adim šibxu 'et 'ozer ha-tabax še-yadu'a be-matkonav ha-yecirati'im
the diners praised ACC helper the-cook that-is.known with-his.recipies the-creative
'The diners praised the help of the cook who is known for his creative recipes.'

39. ha-balaš xakar 'et šaxen ha-kaba'y še-neherag
the-investigator investigated ACC neighbor the-firefighter who-was.killed
'The investigator investigated the neighbor of the firefighter who was killed.'

40. ha-'askan hekim 'et hanhalat ha-kvuca še-nicxa
the-businessman formed ACC management the-team that-won
'The businessman formed the management of the team who won.'

41. ha-'oman šixzer 'et tikrat ha-knesiya še-omedet 'al kce ha-cuk
the-artist restored ACC ceiling the-church that-stands on edge the-cliff
'The artist restored the ceiling of the church that stands on the edge of the cliff.'

42. ha-mišpaxa bikra 'et dod ha-xayal še-yaca lexufat šixrur
the-family visited ACC uncle the-soldier that left for-leave termination
'The family visited the uncle of the soldier that left for a termination leave.'

43. ha-ganan heška' 'et ginat ha-sifriya še-hityabša
the-gardener watered ACC garden the-library that-dried
‘The guard watered the garden of the library that dried.’

44. ha-milioner taram 'et cilum ha-pesel še-hudpas
the-billionaire contributed ACC photo the-statue that-was.printed
‘The billionaire contributed the photo of the statue that was printed.’

45. ha-soxer hexzir 'et masax ha-maxšev še-mašxir
the-merchant returned ACC monitor the-computer that-blackens
‘The merchant returned the monitor of the computer that turns black.’

46. ha-menaka kircefa 'et ricpat ha-mirpeset še-'asuya mi-šayiš nadir
the-cleaner cleaned ACC floor the-balcony that-is.made of-marble rare
‘The cleaner cleaned the floor of the balcony that is made of a rare marble.’

47. ha-kirkas mimen 'et me'alef ha-namer še-hitparsem be-yexolot ha-šlita šelo
the-circus financed ACC tamer the-tiger that-was.known in-abilities the-control of.his
‘The circus financed the tamer of the tiger that was known for his controlling abilities.’

48. ha-taxana acra 'et šidur ha-misxak še-hištabeš
the-station stopped ACC broadcasting the-game that-was.damaged
‘The station stopped the broadcasting of the game that was damaged.’

B-2: Instructions and practice items

Hebrew

שאלון סבירות/קבילות המשפט: הוראות למשתתפ/ת

מחקר דוקטורט זה חוקר את פירושם של משפטים דו-משמעיים. אני בודקת את הגורמים המשפיעים על נטייתם של דוברי השפה העברית לפרש משפטים דו-משמעיים כמו זה שבדוגמה על פי אחד משני הפירושים שלו.

דוגמה: *טלי התיידדה עם מעצבת הספרית שנעלמה*

משפט זה עשוי להתפרש על פי משמעות אחת (בה המעצבת היא שנעלמה) או על פי המשמעות השנייה (הספרית היא שנעלמה). שאלון זה בודק האם שתי האפשרויות סבירות וקבילות מבחינה תחבירית וסמנטית.

במשפטים שבשאלון יוצגו לפניך שתי אפשרויות הפירוש של כל משפט, ומשימתך היא לסמן עבור כל פירוש בנפרד האם הוא מקובל ואפשרי מבחינה תחבירית וסמנטית (5), או האם הוא לא אפשרי/הגיוני כפירוש של המשפט (1). לדוגמה:

טלי התיידדה עם מעצבת הספרית שנעלמה.

המעצבת נעלמה	(5)	4	3	2	1
הספרית נעלמה	(5)	4	3	2	1

שיפוטך לא חייב להיות חד משמעי (כן=5/לא=1), אלא יכול להיות על כל נקודה לאורך הסקאלה: ישנם משפטים שבהחלט יכולים להתפרש על פי שני הפירושים, וכאלו שפירוש אחד (או שניהם) אינו הגיוני כלל מסיבות שונות. להזכירך כי שיפוטך צריך להתבסס על קבילותם/סבירותם של הפירושים.

שאלון זה אינו מוגבל בזמן. כמישהי בעלת רגישות לענייני שפה, נבקשך לשקול היטב את תשובותיך. אם ברצונך להוסיף הערות לצד המשפטים , נשמח לקבלן.

בעמוד הבא יופיע לפניך תרגול קצר. אם המשימה אינה ברורה אנא צרי עמי קשר לפני שתמשיכי בשאלון. אם ההוראות ברורות, תוכלי להמשיך ולתת את שיפוטך על 48 משפטי השאלון. אני מודה לך על עזרתך.

The following is a close translation of the Hebrew original; if the exact wording to the participants is important for some purpose, please consult the original.

PLAUSIBILITY/ACCEPTABILITY QUESTIONNAIRE: INSTRUCTIONS FOR SUBJECTS

This dissertation research examines the interpretation of ambiguous sentences. In particular, it investigates different factors that affect the tendency of Hebrew speakers to interpret ambiguous sentences similar to the one in the example below according to one of its two interpretations.

Example:
Tali became friendly with the designer of the hairdresser who disappeared suddenly.

The sentence above may be interpreted according to one interpretation (in which the designer is the one who disappeared) or according to the other (the hairdresser is the one who disappeared). This questionnaire examines whether the two alternative interpretations are acceptable and plausible syntactically and semantically.

In the sentences in this questionnaire you will be presented with the two interpretations of each sentence, and your task is to mark for each interpretation separately whether it is plausible and acceptable in terms of its semantics/syntax (5) or whether it is impossible/implausible as an interpretation of the sentence (1). For example:

Tali became friendly with the designer of the hairdresser who disappeared.

The designer disappeared.	1	2	3	4	(5)
The hairdresser disappeared.	1	2	3	4	(5)

Your judgment does not have to be binary (yes=5/no=1), but can be marked on any point on the scale: some sentences can definitely be interpreted according to both alternatives, while in other sentences one (or both) of the interpretations are unavailable for different reasons. To remind you, your judgment should be based on the acceptability/plausibility of the sentences.

There is no time limit for completion of this questionnaire. As someone sensitive to language, we asked that you give your considered judgments. I would be happy to receive comments on any of these sentences; please add your comments next to the scale.

In the next page you will have a short practice session. If the task is not clear, please contact me before proceeding. If the instructions are clear, you may continue and give your judgments on the 48 sentences of the questionnaire. Thank you for your help.

CHAPTER 5

PRODUCTION OF DISAMBIGUATED SENTENCES[87]

5.1 Aims

The production study reported in this chapter investigated the prosodic phrasing patterns of the Hebrew RC-attachment ambiguity construction. It examined Hebrew speakers' overt productions of a body of target sentences which had a factorial design crossing construction type (FS/CS) and RC length (short/long RC), as outlined in Chapter 4. In addition, the experimental sentences of the production study were disambiguated, by means of a preceding context, toward either high or low attachment of the RC, adding a third factor to the experimental design: attachment height (high/low).[88] There was reason to expect different phrasing patterns to emerge for sentences with different RC lengths, and/or with different attachment interpretations. Universal constraints on prosodic phrasing and their ranking relative to each other have been argued to determine prosodic structure. Selkirk (1984, 1995, 2000), Truckenbrodt (1999), Ghini (1993), Gee & Grosjean (1983) and others have proposed a set of general constraints on the loci of prosodic phrase boundaries, which are claimed to be universal but to vary between languages with respect to their ranking. Major types of constraints include structural alignment constraints (e.g., align right or left with a syntactic XP); rhythmic constraints (e.g., avoid lengthy unstressed sequences); and length/heaviness constraints (e.g., prefer

[87] Parts of this chapter were reported in Shaked (2007) and presented in Shaked (2006a,b). A preliminary analysis of the perceived phrasing data and attachment data was presented in Shaked et al. (2004).
[88] Attachment to the whole complex nominal is sometimes referred to as N1-attachment, since N1, as the head of the complex nominal, has the property described by the RC. Here, I will refer to it as high attachment.

binary branching in the prosodic hierarchy; prefer equal-length subphrases). It is possible that the relative ranking of the constraints may affect the phrasing of each of the experimental conditions differently.

In particular, Selkirk (2000) offers two constraints which may play a role in determining the phrasing structure of the RC construction under investigation. The first is a syntactic alignment constraint which is parameterized for left or right alignment of prosodic structure and syntactic structure. Considerable evidence indicates that Hebrew, like English, has right alignment; Siloni (2003) notes that in Hebrew "prosodic words are aligned with right edges (right brackets) of lexical words" (p. 488). Selkirk's (2000) *Align$_R$ XP* constraint states that optimally "the right edge of any XP in syntactic structure must be aligned with the right edge of a MaP [Major Phrase; also known as intermediate Phrase –AS] in prosodic structure" (Selkirk, 2000: 232). Thus it is assumed that there is a partial, and asymmetric, alignment between prosodic phrasing and syntactic structure, which targets the ends of phrases in Hebrew as in English. The second relevant constraint is the phonological length constraint *Binary(MaP)* (henceforth: *Bin(MaP)*), which states that optimally "a major phrase consists of just two minor/accentual phrases" (Selkirk, 2000: 244). This length constraint coalesces two distinct length constraints, *BinMin* and *BinMax*. *BinMin* states that "a major phrase must consist of at least two minor/accentual phrases", while *BinMax* states that "a major phrase may consist of at most two minor/accentual phrases" (Selkirk, 2000: 244). The ranking of the length constraints relative to each other and to other types of constraints will determine the language's tendency towards short or longer phrases. I will be adopting Selkirk's constraints in several parts of this dissertation. Selkirk developed these constraints within an optimality

theoretic (OT) framework (Prince & Smolensky, 1993/2004). I discuss these constraints without committing to any specific theoretical framework for representing them, but the evidence presented here may serve as the basis for a formal OT analysis for Hebrew in the future; a preliminary analysis of part of the production data in the spirit of OT is presented in Shaked (2007).

For the RC construction, the interaction of the alignment constraint and length constraints makes different predictions for the phrasing of a long and a short RC interpreted for high and for low attachment. High- and low-attaching RCs have different syntactic structure, as Figure 5-1 schematically illustrates.

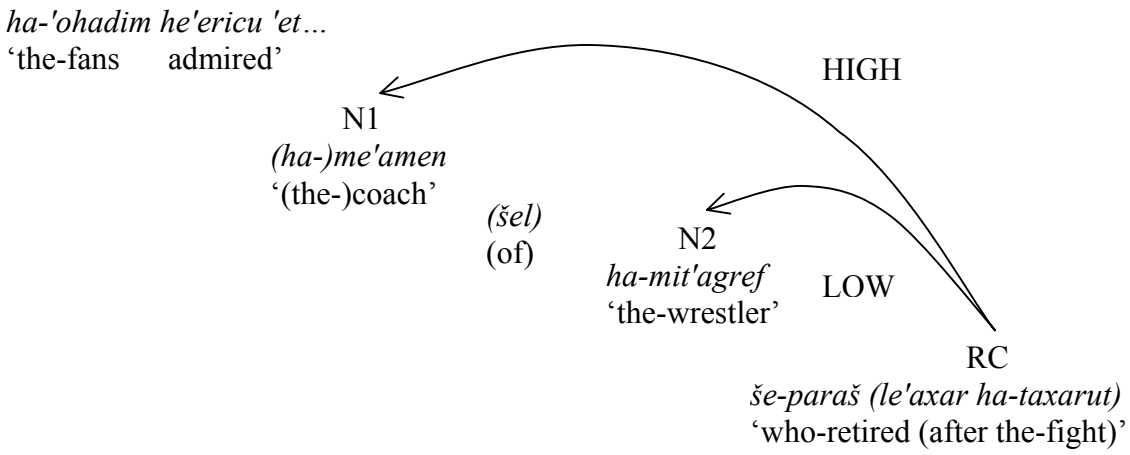


Figure 5-1: Schematic illustration of high and low attachment of the RC (of different lengths) to a preceding complex nominal.

Figure 5-2 and 5-3 provide a detailed surface structure for FS and for CS, respectively, based on the syntactic analyses suggested by Siloni (1996, 1997; see also Chapter 2, §2.5). Siloni's analysis is neutral on the issue of the category of *šel* ('of') and in particular whether it is or it is not a preposition. Since Siloni's analysis was developed out of

Ritter's (1988), who does treat *šel* as a preposition, for concreteness I've made a working assumption here that *šel* is a preposition. Following Siloni's proposal that *šel* assigns Case to its complement, it is assumed here that DP_2 is the complement of *šel*. In addition, I assume that restrictive RCs attach as adjuncts within NPs (in contrast to the association of non-restrictive RCs within DPs).[89] The potential placement of RCs associated with N1 and N2 is marked by dotted syntactic branching, to express their optional status in the structure.[90]

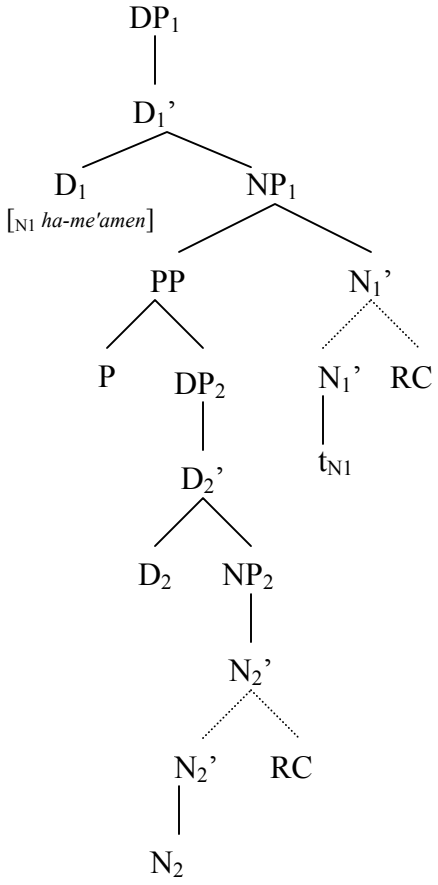


Figure 5-2: Derived structures of FS, with high and low RC associations.

[89] It is not entirely clear whether modifying adjectives could be assumed to occupy a similar position to RCs, since without further movement operations the wrong word order would follow. Specifically, an adjective modifying N1 would surface in the position following the whole complex nominal whereas in fact it should immediately follow N1. Siloni assumes that adjectives are left adjoined to NPs, with movement of the N to a preceding position, as discussed in Chapter 2, §2.5.

[90] Note that the structures presented here explicitly show the traces of movement operations; earlier work (Shaked, 2007) ignored this aspect of the analysis.

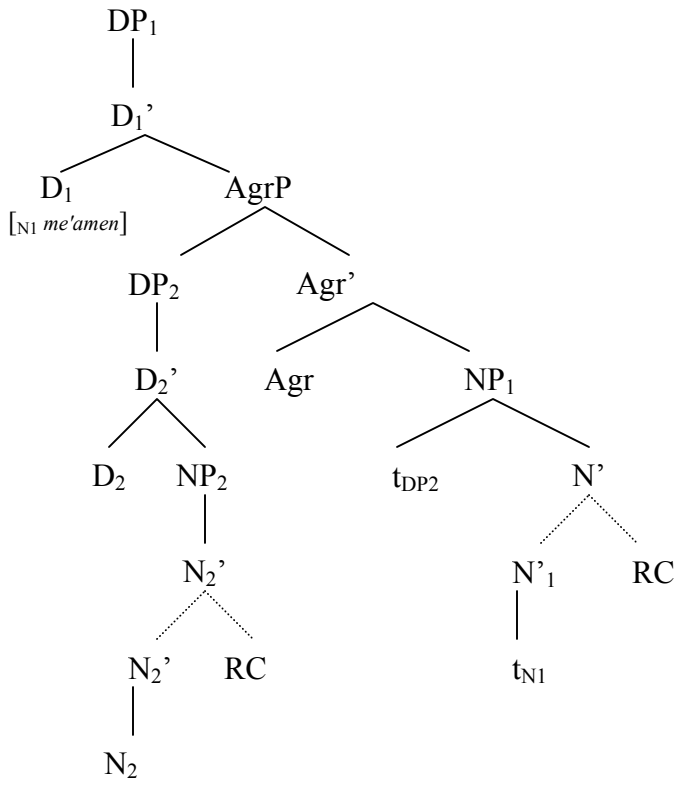


Figure 5-3: Derived structures of CS, with high and low RC associations.

In the high-attaching configuration, the RC modifies the whole complex nominal and there are several maximal projections preceding the RC. (For FS sentences, the maximal projections PP, DP_2, and NP_2, precede the RC; for CS sentences, the maximal projections DP_2 and NP_2 precede the RC). There is no indication here that there is an accumulating effect of the number of brackets (i.e., that having an extra bracket in the FS affects the likelihood of a prosodic break in that position; but see Schafer, 1997; Frazier and Clifton, 1998; and Fodor, 2002). As the data below show, however, there might be a ceiling effect. The alignment constraint predicts a prosodic break in this position, i.e., following the complex nominal and preceding the RC, to align with the discontinuity in the syntactic tree. In the low-attaching configuration, by contrast, the RC modifies the

second noun only. There is no right edge of a maximal projection intervening between the N2 and the RC when the RC attaches low, since the RC is attached within NP_2 (the structure of NP_2 is similar for both FS and CS sentences). Thus the alignment constraint does not predict a prosodic break to occur after N2 in the low attachment configuration.

Setting aside alignment constraints, length constraints in and of themselves may differentially affect the phrasing of sentences with varying RC lengths. A long RC, heavy enough to form a prosodic phrase on its own, is likely to be phrased separately from the complex NP, by placement of a prosodic boundary between N2 and the RC. A short RC, on the other hand, is light and therefore is predicted by the prosodic length constraints to be included in a prosodic unit with its neighbor N2. The production experiment thus considers the relative contributions of different kinds of prosodic constraints on phrasing decisions, looking specifically for effects of RC length and forced attachment height.

In addition, there was reason to expect differences in the phrasing patterns of sentences containing the FS and CS. Recall from Chapter 2 that the CS is standardly described as constituting a single phonological word. As such, no break is predicted to occur between N1 and N2. The FS, on the other hand, contains two separate nouns, which constitute two separate phonological words; therefore, a phrasal break can and might (but also might not) occur between these nouns.[91] The production experiment explores the phrasing strategies used for sentences containing these constructions. The study also provides experimental data that may be used as support for the status of the CS as a single phonological unit.

[91] It is not entirely clear what the right unit is for calculating length constraints. Selkirk's length constraints refer to minor/accentual phrases. In the spirit of Selkirk's formulation, I am counting units which each contains one accent. In this discussion I am referring to these as prosodic words.

The outcomes of the production study thus provide evidence of the variety of phrasing patterns used by Hebrew speakers in conveying the meanings of these RC constructions. It explores whether different phrasing patterns are associated with the different experimental conditions under investigation (a factorial combination of FS/CS; short/long RC and high/low attachment interpretation). And for each condition it reveals the preferred phrasing pattern and the relative proportions of alternative patterns. In these ways it may enhance our knowledge of sentence-level prosodic phrasing in Hebrew, with possible applications to the study of other utterance types as well, thus making a contribution to a general characterization of the prosodic structure of Hebrew.

The production data will also serve as an important background for the listening experiment (Chapter 6) and the silent reading experiment (Chapter 7), which employ fully ambiguous target items. The task employed in this production experiment yields semi-spontaneous productions that are contextually disambiguated for high or for low attachment. Thus, the target sentences in this experiment were produced with one unambiguous meaning (and hence syntactic structure) in mind. The array of phrasing patterns exhibited for each particular meaning can shed light on what expectations about prosody hearers bring to the listening task, and what prosodic patterns readers are likely to impose on written sentences, whose prosody is not signaled in the stimulus.

Especially useful is the fact that it can help to clarify the notion of default prosody, which is central to the Implicit Prosody Hypothesis (IPH) for silent reading (see Chapter 1) but which can be difficult to pin down for ambiguous items. For disambiguated items as in the production experiment, if a given sentence type exhibits a consistent phrasing pattern regardless of the attachment height, it could be concluded that

that phrasing pattern is the default prosodic structure of that sentence type, and the IPH then makes predictions about attachment preferences in silent reading. If, on the other hand, different prosodic patterns occur equally robustly within the same condition or across the same sentence type in different attachment interpretations, then this suggests that there is no stable default prosodic structure for that sentence type and hence that the IPH is not operative, i.e., implicit default prosody is not what is guiding attachment preferences. Thus, exploring differences and similarities in the range of prosodic patterns associated with the different experimental conditions in the production study with disambiguated RC sentences may prove revealing in exploring how perceivers process the ambiguous RC construction in Hebrew. In sum: there are multiple motivations for the production experiment, since it can be informative of both the prosody-syntax interface in the 'competence' grammar of Hebrew, and also 'performance' aspects of speaking, listening and reading, and how they relate to each other.

5.2 Method

5.2.1 Materials

The materials for the production study were built around 8 basic items. These were a representative subset of the 24 target items constructed for the other experiments reported in this dissertation and therefore subject to the criteria described in Chapter 4. Performance variability in production studies being relatively low by comparison with comprehension studies, smaller item sets were expected to provide sufficient power to evaluate the hypotheses at issue, and this was borne out in the results presented below. The full list of experimental materials is presented in Appendix A-2 to this dissertation.

All target items had a complex nominal (either an FS or a CS), followed by a relative clause (either a short or a long variant). The 8 items chosen for this experiment had an equal number of animate and inanimate complex nominals, similar to the ratio in the full 24 items investigated in this dissertation. There were no filler items in this experiment.[92]

One of the phonological consequences of the CS constituting a single phonological unit is that N1 in the CS can (though it does not necessarily) undergo morphological alternations as a result of stress reduction (see discussion in Chapter 2, §2.3). Although the meaning of the CS is unaffected, this phonological alternation can make an acoustic comparison of the CS and the FS problematic. However, such comparison was necessary for the analysis of data from the production study, so a strict criterion for selection of items for the production experiment was imposed: N1 should have a similar phonetic form in the CS and in the FS. By this I mean that no phonological alternation such as vowel deletion and/or final consonant deletion takes place in the CS. In fact, N1 may not be completely identical in the two complex nominals, since it undergoes stress reduction in its bound form in the CS, as described in Chapter 2, while there is no stress reduction in its free form in the FS nominal. This, however, does not seem to have affected the results, e.g., the high attachment data shows that N1 had similar f_0 and duration means in CS and FS, and there was no effect of construction type in the instrumental analysis (see §5.3.2 below for details).

[92] It was deemed unnecessary to disguise the constructions under investigation by using filler items in this experiment because the goal was to examine differences in the phonological/acoustic properties of speakers' productions of the experimental sentences. Although subjects may have noticed aspects of the factorial design (e.g. that the sentences are ambiguous but the procedure disambiguates the RC as modifying only one of the two nouns), it is unlikely that they were aware that the focus of the study was the prosodic patterns for the different sentence types.

Although the target sentences themselves were fully ambiguous structurally, they were contextually disambiguated for either high or low attachment. The protocol used in the production experiment showed a simple phrase (e.g., 'The fans admired the coach of the wrestler'), followed by a question and an answer (e.g., 'Which wrestler?'; 'The wrestler who retired'). Together, the three phrases constituted a clearly disambiguated context, to be used by participants to create a complex utterance with one unambiguous meaning (e.g., an utterance of 'The fans admired the coach of the wrestler who retired' where the wrestler is the one retired). This is the 'Post-to-Times' disambiguation technique originated by Bradley et al. (2003), which will be described more fully in §5.2.3. Since the target sentences were presented to participants with a clear unambiguous meaning, their production of the sentences can be taken as representative of the prosodic patterns that are associated with this meaning.

5.2.2 Design

The production study had 3 fully crossed factors. It included the same two factors repeated in all three experiments: Construction type (FS or CS) and RC length (short RC or long RC). An additional binary factor in this experiment was height of attachment: the utterance elicitation context either forced the interpretation with the RC attaching high ("forced high"), or with the RC attaching low ("forced low"). When fully crossed, the three binary factors created 8 conditions per item: forced high attachment interpretation of FS and CS with both short and long RC and forced low attachment interpretation of the two complex nominals (FS and CS) in the two RC lengths (short and long).

While each participant was exposed to both interpretations (forced high-attaching and forced low-attaching) of each sentence during the course of the experiment, these two meanings were not mixed. The experiment was designed such that the list of sentences was divided into two blocks: one block presented the 4 conditions of each item (created by the construction type and RC length design) disambiguated for a low-attaching interpretation ('low' block), while the other block presented the same quadruple disambiguated for a high-attaching interpretation ('high' block). Participants completed one block of experimental items (one booklet of items with a consistent attachment interpretation throughout), and after a few minutes break they proceded to the second block of experimental items (a second booklet, with the opposite attachment interpretation). The reason why participants were presented with only one type of interpretation throughout each block was in order to guard against confusion about the interpretation intended, and to minimize any use of a deliberately disambiguating (e.g., contrastive) prosody to convey the difference in attachment choice.[93]

Half of the participants saw the 'low' block first, followed by the 'high' block; for the other half, the two blocks were presented in reverse order. This created two versions for the questionnaire, varying in order of presentation of the two blocks. Items in each block were ordered according to an identical pseudo-randomization. The pseudo-randomization ensured that no two versions of a given item were presented one

[93] The data reported in this dissertation include only the first block encountered by each participant (for half of the participants, a forced-high attachment interpretation and for the other half a forced-low attachment interpretation), thus making attachment height a between-subjects factor. Data from the second block of the questionnaire are not reported here, for reasons as mentioned above. Some participants (in each group) commented at the end of the experiment that the second part of the experiment (i.e., that with the opposite attachment choice) was more difficult for them. It was also the author's impression that most participants had more trouble with performing the task in the second part, as evidenced by more frequent hesitations and the consequent need to repeat the recordings than in the first part. It seems that participants found it difficult to switch between one interpretation and the other; and in coping with this, some responses were characterized by an extreme use of contrastive stress.

immediately after the other, and that there were no more than 3 consecutive sentences with the same RC length.

There were a total of 32 target sentences in each block, constructed as 8 quadruples and defined by FS/CS and short/long RC. In addition to the 32 target items, the experiment included 10 additional sentences that were similar in structure to target sentences but which were not incorporated into the analysis. These filler sentences were designated as 'block protectors', which were selected from among the larger set of 24 target items constructed for the dissertation study, but which were not part of the subset of 8 experimental sentences chosen for the production study. They were inserted as covert practice items at the beginning of a block, in order to protect target items against an unnatural reading, and at the end of a block as protectors against a more sloppy reading due to anticipation of the end of the experiment. Each block began with two block protector sentences and ended with 4 block protector sentences.

The experiment was self paced, and was designed such that there were two designated (optional) breaks within each block of the questionnaire, to allow participants some rest. These two break positions were followed by 2 block protector sentences, in order to allow participants some warm-up, similar to the use of the block protectors at the beginning of the block. Thus in addition to the 32 experimental items, each of the 2 blocks had 10 block protector items. The experiment began with an instructions phase which included two illustrations of how to perform the task, followed by 3 practice items. Thus, in the course of the experiment each participant encountered a total of 89 sentences: 64 experimental sentences (32 in each block), 20 'block protector' items (10 in

each block), 2 illustrations and 3 practice items. The experiment took approximately 50 minutes to complete, including instructions and practice.

5.2.3 Procedure

The experiment employed an elicitation technique developed by Bradley et al. (2003), called the 'Post to Times' protocol.[94] The list of items was presented in a booklet. Each page contained a single item. For each item, participants saw three written phrases: a statement, a question about the statement, and an answer to the question, printed on three separate lines in the middle of the page. The three written phrases together unambiguously indicated the meaning of the intended complex sentence: The initial statement was a simple sentence which presented the complex nominal; the question and answer that followed determined the attachment of the RC (to the first noun or to the second noun), thus restricting the interpretation of the RC. Participants were instructed to read the three lines of text, first silently (for comprehension) and then aloud, and to mentally form a complex sentence by combining the statement and the answer into a single sentence containing the RC and then utter it aloud (into a voice recorder). It should be noted that this task of combining simple sentences into a complex sentence is not an infrequent task in grade school, so there is no reason to think that the participants found it unexpected or unnatural.

[94] The name of the protocol, the 'Post to Times', originated from the idea of taking simple phrases or sentences (a type characteristic of the *New York Post* newspaper) and combining them to form a complex sentence (more typical to the *New York Times* newspaper).

This protocol was designed to ensure that participants' utterances reflect a specific meaning,[95] as determined by the experimenter, but without any change in the lexical content of the sentence: the sentences produced had exactly the same word sequence to be analyzed acoustically, which is an advantage. There are other purely contextual approaches to disambiguation with the same advantage, such as Hirshberg & Avesani (1997), who present a paragraph in which the target sentence is embedded. But Bradley et al. (2003) maintain that such procedures make it more likely that participants would simply read the sentences aloud, while their approach is more like natural sentence production, even if it cannot be fully so. The task of the Post-to-Times protocol kept participants active (rather than passive) and they were free to impose any prosodic phrasing they wished. Importantly, in the instructions and the practice session prior to the experiment, their attention was not drawn to the prosody of their utterances. Also, no attention was drawn to the syntactic and semantic ambiguity of the elicited sentence.[96] An example of the protocol for a sentence (containing an FS with a short RC) is presented in (1a) for a forced-high attachment interpretation and in (1b) for the forced-low attachment interpretation; (1c) shows the structurally ambiguous target sentence that was to be

[95] Previous studies of the RC construction in other languages appear to have assumed that perceivers interpret the RC restrictively (despite the absence of any special discourse that would ensure this). Related to this is that it is also unclear in some cases whether the boundary preceding the RC was an intermediate phrase (ip) boundary or an intonational phrase (IP) boundary. I used the elicitation protocol described above in order to resolve these uncertainties. It elicits utterances in which the RC clearly has a restrictive interpretation. Informal observations suggest that a restrictive RC in Hebrew is typically preceded by an ip rather than an IP boundary. Indeed, ip boundaries were observed in most of the data of this experiment; in a few cases, an IP boundary was observed nonetheless.

[96] In light of the debate in the current literature (e.g., Allbritton et al., 1996; Fox Tree & Meijer, 2000; Schafer et al., 2000; Snedeker & Trueswell, 2003; Kraljic & Brennen, 2005; Schafer, Speer, & Warren, 2005) on the extent to which participants do or do not deliberately disambiguate while performing production tasks with ambiguous materials, participants in this study were presented with a final debriefing questionnaire querying whether they had noticed the ambiguity. Several participants had indeed noticed the ambiguity, but unfortunately, since this was an end-of-experiment questionnaire, it did not distinguish between the perception of ambiguity after block 1 (whose data are reported here; see above) compared to after block 2, which could well have triggered awareness of the existence of an alternative interpretation (see footnote 90).

created by participants, whether interpreted high or low. Note that the same target sentence was created for both variants in (1), although the prosodic phrasing assigned to each production by the participant was expected to be different, as the results below confirmed.

(1) a. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'argef.
the-fans admired the-coach of the-wrestler

eize me'amen ?
which coach

ha- me'amen še-paraš.
the-coach who-retired

b. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'argef.
the-fans admired the-coach of the-wrestler

eize mit'agref ?
which wrestler

ha-mit'agref še-paraš.
the-wrestler who-retired

c. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'argref še-paraš.
the-fans admired the-coach of the-wrestler who-retired

Participants were tested individually in a quiet room within the Linguistics Department of Tel Aviv University. An instruction session was provided prior to the beginning of the task, demonstrating how the two sentences were to be structurally combined. The instructions are presented in Appendix B-1. The instructions emphasized that participants needed to provide a recording that is fluent, natural and coherent, and therefore they should repeat each item that was unsatisfactory (i.e., included hesitations, mispronunciations of words or unnatural sounding). Participants were also told to expect a given sentence to appear several times, with slight modifications, so they should not

assume it would be identical to a previous similar sentence; this prepared them for the repetition of items in the 8 different conditions in the course of the experiment (4 repetitions of each item per block).

A practice session then followed, in which participants performed the task on three practice items, for which they received feedback if necessary (specifying what constituted a hesitation or an unnatural reading, and how to perform the task of combining the phrases into a complex sentence correctly). They were encouraged to ask questions before proceeding to the experiment itself, but once they began the experiment, they were instructed to stop only in the two designated positions, clearly marked in the experiment booklet, if they desired a break. Participants' utterances, including the statement, question and answer that preceded the target sentence, were recorded using a Sony ICD-BP150 digital voice recorder, which was connected to a head-mounted microphone to eliminate amplitude variation due to head movement.

5.2.4 Participants

Twenty-four Hebrew speakers participated in the experiment successfully: 7 males and 17 females. 12 encountered the 'high' block first and 12 encountered the 'low' block first. All participants were naïve with respect to the purpose of the experiment. They were students at Tel Aviv University in their twenties or early thirties. They were all native speakers of Hebrew, and Hebrew was the primary language they used in their daily life.

Participants filled out a language background questionnaire before they started the experiment, which asked about their language use in order to establish that they were

indeed Hebrew native speakers, and that there was no other reason to exclude their responses from the data analysis. A copy of the background questionnaire is provided in Appendix C. Participants were volunteers, and received a small payment (30 NIS, equivalent to $6 at that time, typical for this kind of participation in Israel) before the experiment began.

5.2.5 Data measurement and assessment

The utterances produced using the elicitation protocol provided raw data which were used to establish the phrasing patterns of Hebrew native speakers. Two distinct methods were applied for establishing the phrasing patterns of the target construction: one relied on human perception of the position of prosodic breaks in the sentence, using a coding system similar to that of the Tones and Break Indices (ToBI) system (Beckman and Hirschberg, 1994); the other employed acoustic analyses which examined f_0 and duration data for N1 and N2. The specific procedures for establishing the phrasing patterns by each method are discussed separately below.

5.2.5.1 Perceived phrasing data

The perceived phrasing analysis looked for different phrasing patterns associated with the target construction, and effects of the design factors on the phrasing patterns used. Based on the literature on Hebrew prosody presented in Chapter 3, a preliminary prosodic transcription system, based on the Tones and Break Indices (ToBI) transcription system, was developed. A complete ToBI transcription system, comparable to that developed for several languages based on the original work by Beckman & Hirschberg

(1994) on English (see Jun, 2005, for a description of the ToBI systems for 11 languages), has not yet been developed for Hebrew, so the coding system employed in this study is necessarily tentative at present.

The ToBI notational system for transcription of intonation (Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005) assumes the Autosegmental-Metrical (AM) model of intonational phonology (Pierrehumbert, 1980; Ladd, 1996). Though the system was standardized in the first instance for English, the inventory of tonal elements from which it may draw can be adapted to transcribe intonation in other languages, hence J_ToBI for Japanese, G_ToBI for German, and so on. ToBI transcriptions, like the intonational phonologies from which they derive, adopt a tone-sequence approach to distinguish different utterance intonations. That is, they draw from language-specific inventories of simple tones (H, L) or tone-pairs (e.g., H*+L) to represent categorically distinct pitch accents that are associated with elements in the segmental string; they also assign boundary tones (H, L) to the edges of hierarchically organized prosodic domains. To provide phonetic evidence relevant to ongoing intonational research, ToBI transcriptions also designate 'break indices' for each pair of surface-adjacent words. These range in value from 0 for cases of maximal integration, e.g., between a word and an associated clitic element, to 4 for cases of maximal non-integration, i.e., a full intonational phrase break. See Beckman, Hirschberg & Shattuck-Hufnagel (2005) for a more detailed account of the development of the ToBI framework.

The ToBI system for Modern American English (Beckman, Hirschberg & Shattuck-Hufnagel, 2005) assumes 3 types of tones: pitch accents (placed on the stressed

syllable of each prosodic word), phrase accents (obligatorily placed at the end of an intermediate phrase) and boundary tones (obligatorily placed at the end of an intonational phrase). It further assumes the following break indices, indicating the perceived degree of boundary strength: 0 (very close inter-word juncture), 1 (ordinary phrase-internal word end), 3 (intermediate phrase end, with phrase accent) and 4 (intonational phrase end, with boundary tone);[97] each intonational phrase contains at least one intermediate phrase, and each intermediate phrase contains at least one pitch accented word.

As discussed in Chapter 3, §3.2, the literature on Hebrew prosody generally distinguishes between two types of phrase boundary in Hebrew: a "breath group" boundary (comparable to Break Index 4 (BI=4) in the ToBI system, as argued in Chapter 3) and a "sense group" boundary (corresponding to Break Index 3 (BI=3) in the ToBI system). Following the work of Beckman and her colleagues, it is possible to postulate for Hebrew the break indices described in Table 5-1.[98]

[97] Break Index 2 is a disjuncture which is perceived as 1, but has an unexpected tone marker or lengthening more suitable for break index 3 or 4, though without the expected phrase accent or boundary tone. In addition to the BI markings, several diacritics provide additional information, e.g., '?' for uncertainty.

[98] I would like to thank Mary Beckman for her advice on Hebrew prosody and for suggesting the possible types of break indices available for Hebrew. All errors of judgment in this regard are mine.

Table 5-1: Proposed break indices for Hebrew and their definitions, modeled on the ToBI system of Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005).[99]

Break index	Definition
BI=0	Phonetically marked boundaries internal to the clitic group (e.g., *ha-* 'the')
BI=1	Clitic-group boundaries which are not prosodic word boundaries (e.g., boundaries internal to compounds and construct state nominals).
BI=2	Minor/accentual phrase boundaries: Word boundaries marked by pitch accents.
BI=3	Major/intermediate phrase boundaries: phrase boundaries marked by a single rising phrase accent and optional pausing/lengthening.
BI=4	Full intonation phrase boundaries, marked by a falling boundary tone and obligatory pausing/lengthening.

For the purposes of this study, it was of primary interest to identify intermediate phrase boundaries, corresponding to BI=3, since it is these that, according to Selkirk's hypothesis, affect prosodic phrasing (see discussion above), and according to native speakers' intuitive judgments, these phrase boundaries differentiate the alternative possible phrasings of the complex NP construction with a restrictive RC. A very few sentence-internal, full intonation phrase boundaries occurred in the utterances produced by participants (.02 of utterances); these were characterized by a falling boundary tone.[100]

Intermediate phrase boundaries in Hebrew are associated with three prosodic cues (Enoch & Kaplan, 1969; Laufer, 1987, 1996): (a) a rising boundary tone on the accented syllable and then an f_0 resetting after the boundary, (b) pre-boundary lengthening of the

[99] See Chapter 3 for discussion of clitic group boundaries corresponding to BI=0 and BI=1.
[100] A full intonation phrase boundary (IP) is typically characterized by (obligatory and sometimes extensive) lengthening, and a falling tone at the end of the sentence or before a non-restrictive RC sentence internally. In constructions with coordinate clauses or an initial adverbial clause, the sentence-internal IP boundary seems to be marked by either a preceding continuation rise or a falling-rising tone, accompanied by a distinct (obligatory) lengthening.

last syllable of the phrase and a faster speech rate at the beginning of the next intermediate phrase, and/or (c) a pause (silence). Among these, there is consensus among researchers on Hebrew that the most reliable cue for an intermediate phrase boundary is a change in pitch height, though it may be accompanied by pre-boundary lengthening and occasionally by optional pausing as well (see Chapter 3, §3.2 for more details).

The default lexical stress in Hebrew is assigned to the final syllable of a word (Chayen, 1973; Bolozky, 2000; Graf, 2000; Ussishkin, 2000; among others), and is realized by increased vowel length and high tone (Becker, 2002). A possible objection can be raised as to whether perceivers can distinguish between word-level pitch accent and phrasal-level boundary tone, both realized by similar prosodic cues (changes in the pitch contour of a stressed syllable). As noted in the guidelines for ToBI labeling (Beckman & Ayers Elam, 1994:10), pitch accents and phrase boundaries differ in their predictability: Stress patterns (stressed and unstressed syllables) within each word are largely predictable, since they are fairly fixed and can be seen as an inherent feature of the word. Phrasal stress patterns, on the other hand, are not as predictable, and may vary in different utterances of the same sentence. It seems plausible that perceivers distinguish stress patterns above the word level from within-word stress patterns on the basis of sharper rises in fundamental frequency, which are unpredicted by the lexical items themselves. Laufer (1987) maintains that indeed perceivers do discriminate on this basis; he notes that "the native ear can notice a small change in the tone level in order to identify that the syllable is stressed, but such a small change is not perceived as a change in tone in terms of prosody" (Laufer, 1987:47; translation by AS). It is therefore expected, following Laufer (1987), that the trained judges who marked perceived

phrasing judgments on the production data are able to make the distinction between the different boundary types. The acoustic data reported below support this assumption. It can also be seen as supported by data presented in Amir, Almogi & Gal (2004), discussed in Chapter 3, §3.1.

Based on cues for intermediate phrase boundaries proposed in the Hebrew phonological literature, the notation system presented in Table 5-2 was developed for use in this study. Definitions for each notation are based on Beckman & Hirschberg's (1994) definitions, with my adaptations as noted in the table.

Table 5-2: A notation system for identifying Hebrew intermediate phrase boundaries. Definitions are based on (Beckman & Hirschberg, 1994; Beckman & Ayers-Elam, 1997; Beckman, Hirschberg & Shattuck-Hufnagel, 2005), with my adaptations.[101]

Notation	Definition
3	Pausing/lengthening indicating an intermediate phrase boundary
3-	Uncertainty about strength of pause/lengthening (whether it is BI=2 or BI=3)
3p	Prolongation suggesting hesitation, or markedly long pauses
H-	Lexical stress (‘peak accent’) + rising boundary accent on final syllable
H-?	Uncertainty over the tonal value on final syllable (whether it is only a pitch accent (H*) or also a phrasal accent (H*H-))
L-	Lexical stress (‘low accent’) + falling boundary tone on final syllable
L-?	Uncertainty over the tonal value on final syllable (whether it is only a pitch accent (L*) or also a phrasal tone (L*L-))

[101] Table 5-2 is a notation system designed to transcribe intermediate phrase boundaries, therefore it focuses on break index type 3 in terms of lengthening/pausing and in terms of phrase boundary accents. Since it is assumed that these are intermediate phrase boundaries, it necessarily is the case that a phrase boundary is accompanied by a pitch accent placed on the stressed syllable of the phrase-final word. I do not specify here the different pitch accents available in Hebrew, although some related discussion is presented in Chapter 3.

The L* pattern was not expected to be observed except for an occasional IP boundary in the data (see footnote 96). This coding system was employed by two native speakers of Hebrew, who found it natural to use and who agreed largely in their judgments; see below. The appropriateness of the coding system is further supported by the close parallelism between the phrasing data and the acoustic data reported in §5.3.

The procedure for coding the perceived phrasing patterns was as follows: Two trained judges identified prosodic boundaries anywhere in the sentence, using the notations in Table 5-2. The first judge was the author of this dissertation, and the second was an MA student in Speech and Hearing Sciences at the City University of New York who was naïve with respect to the procedures and goals of the experiment. She was trained in use of the coding system by the author. Both judges are native speakers of Hebrew. They each listened to target sentence productions only and did not hear the context that disambiguated for high or low attachment, thus ensuring impartial judgments.[102] They marked tonal and pausing/lengthening cues for intermediate phrase boundaries by searching for each type of cue in a separate pass through the sentence. After doing a complete coding of all the utterances, each judge then went over her own coding a second time, making a forced decision as to whether or not an intermediate phrase boundary was present (eliminating the notations that indicated uncertainty, i.e., 3-,

[102] Judges marked for each participant the four utterances of an item as a group, in order to allow a comparison of the pitch track of a participant in the four different sentence types. Different speakers have different pitch contours and speech rates. Thus a speaker whose fundamental frequency is relatively low has lower rises relative to a speaker with higher pitch contour. In the same way, a speaker whose speech rate is very fast has less prominent pauses and/or final syllable lengthening relative to a slower paced speaker. Therefore, for each participant and for every item, the utterance was compared to the three other versions of the sentence. In this way we extracted perceived judgments that are relative to the acoustic characteristics of each participant. Judges first listened to the group of 4 utterances of item A across all participants, then to item B across all participants, and so forth.

H*H-? and L*L-?).[103] A comparison of this forced-decision coding by the two judges revealed an overall agreement of 97.7%. The two judges deliberated and together were able to come to an agreement on the remaining cases.

A statistical analysis of the perceived phrasing data provided two kinds of results, reported in §5.3 below: first, it provided the range of prosodic phrasing patterns which Hebrew speakers used in expressing the target construction in its different versions. Second, it provided the preferred phrasing pattern for each of the experimental conditions, and the frequencies with which alternative phrasing patterns are used. Before presenting the results, I will first describe the procedure for assessment of the acoustic data.

5.2.5.2 Acoustic data

As noted in §5.2.5.1 above and in Chapter 3, intermediate phrase boundaries in Hebrew are realized by a high boundary tone, and pre-boundary lengthening with optional pausing. The acoustic markers signalling the end of an intermediate phrase are thus a rising fundamental frequency (f_0) and longer durations for preceding syllables. As in the perceived phrasing analysis, the acoustic analysis looked to distinguish alternative prosodies used for the target construction: the phrasing *N1)(N2 RC* puts N1 in phrase-final position; *N1 N2)(RC* does the same for N2; and the prosody imposing no boundary internal to the complex nominal, *N1 N2 RC*, makes neither of the nouns phrase-final. Acoustic evidence distinguishing the different experimental conditions therefore arises in

[103] Judges characterized a break with notations as to the prominence of the signal, i.e., whether they perceive a prominent pause and lengthening (3) or a suggestive pause and lengthening (3-), and whether the changes in the pitch track were significantly or slightly noticeable (e.g., H*H- or H*H-?, respectively). In order for a marking to be considered a clear prosodic break to be included in the statistical analysis, at least one of the signals of a break (i.e. lengthening, pausing or resetting of pitch track) needed to be prominent.

f_0 and duration data for N1 and N2: an f_0 rise and/or longer duration on N1 would indicate a prosodic break after N1, while an f_0 rise and/or extended duration on N2 would point to a prosodic break after N2. No acoustic effect in either N1 or N2 would indicate that no prosodic break occurred after either of the nouns.

Acoustic measures cannot be interpreted absolutely since they are affected by item-specific phonetic content.[104] For this reason, the analyses that follow cannot compare N1 values to N2 values; rather, they must compare f_0 and duration values for each of N1 and N2 separately, across utterances of the same basic item elicited under different experimental conditions. For example, the f_0 value of N1 in the low-attaching short RC condition would be compared to the f_0 value of N1 in the high-attaching short RC condition, and in the low-attaching long RC condition; differences in f_0 values of N1 between the conditions would suggest that a prosodic break has occurred in one condition, but not in the other.

The instrumental analysis extracted measurements of 6 regions, which were comparable across the 8 conditions of the experimental design. Acoustic data are presented for the N1 and N2 regions (regions 3 and 5 below); data for the remaining regions were extracted, but were not used for the instrumental analysis. Each waveform

---

[104] Duration necessarily varies with a word's syllabic and segmental composition. f_0 varies intrinsically as a function of vowel height, e.g., Most, Amir & Tobin (2000) estimate that the Hebrew high vowels /i/ and /u/ have f_0 values as much as 20 Hz greater than the low vowel /a/, for some speakers. Consider, for example, bisyllabic nouns *yadit* 'handle' and *dira* 'apartment', in which a sequence of a low and a high vowel or vice versa would decrease and increase, respectively, the f_0 rise observed over the entire word, for purely phonetic rather than phonological reasons. It should be noted that the f_0 contours discussed below as indicative of prosodic boundaries exhibit a rise beyond that attributable to lexical stress. This is evidenced in a comparison of f_0 measurements of a N1 (or of N2) across different conditions, which shows that the measurement in one condition is greater compared to its measurement in the other conditions, suggestive of an additional rise in tone beyond the lexical stress; see below. Since vowel quality was not matched across N1 and N2, the lexical stress is modulated differently in N1 and N2 by the inherent f_0 vowel quality biases, as explained above, therefore each of N1 and N2 are compared against its own occurrence in the different conditions, while measurements of N1 and N2 are not compared.

was segmented using acoustic analysis software (Speech Analyzer, v. 2.5). The segmentation plan is presented in table 5-3.

Table 5-3: Segmentation of waveform into 6 regions comparable between the 4 experimental conditions.

Region	Segment	FS	CS	Transliteration
1	Subject	ha-'ohadim	ha-'ohadim	the-fans
2	verb (definite article)	he'ericu 'et ha-	he'ericu 'et	admired (the)
3	N1	**me'amen**	**me'amen**	coach
4	(šel) definite article	šel ha-	ha-	(of) the
5	N2	**mit'agref**	**mit'agref**	wrestler
6	RC	še-paraš (le'axar ha-taxarut)	še-paraš (le'axar ha-taxarut	who-retired (after the-fight)

This segmentation plan isolated the defined N1 and N2 regions, which enabled the extraction of f_0 and duration data of these segments. Figure 5-4 presents an illustration of a segmented target item, for the case of an FS with a short RC. The segmentation plan ensured the consistency of decisions about the limits of each region: for any basic item, landmark acoustic criteria were identified in advance, and applied across all variants of that item, and for all speakers. Note that the segmentation plan ensured that N1 and N2 consist of the same phonological material in both the CS and the FS. This allows for a comparison of measurements of these regions across sentence types despite differences in the surrounding phonological content.[105] Each segment ended at the beginning of the

[105] Differences in segments surrounding N1 and N2 are the result of the absence of a definite marker on N1 in the CS, and the presence of the element *šel* in the FS. The N1 and N2 regions were stripped of any

phonetic content of the next segment, such that any pausing or lengthening after a word in the sentence was captured within its segment, being attributed to the closing of a prosodic phrase. Thus, if it was found that for two items that differ minimally by a manipulated factor, a segment was significantly longer or had a significant rise in f_0 track in one sentence type relative to the other, it was concluded that a prosodic break occurred after the word in the former sentence type. More specifically, if it was found, for example, that the N1 region was significantly longer and/or had higher pitch estimates for an FS with a short RC than with a long RC, then it would be concluded that the longer/higher pitch N1 segment contains a phrasal break following the noun. Under the prediction that prosody would be affected by the attachment interpretation of the sentence as well as by the interface between the grammar and prosodic rules of Hebrew, it was expected that the incidence of a phrasal break between the two nouns (i.e., following N1) or following N2 might vary across the interpretations, when a speaker produced a contextually disambiguated sentence.

---

elements except for the nouns themselves (including the definite articles), in order to be able to have a fully comparable instrumental measurements of these regions.

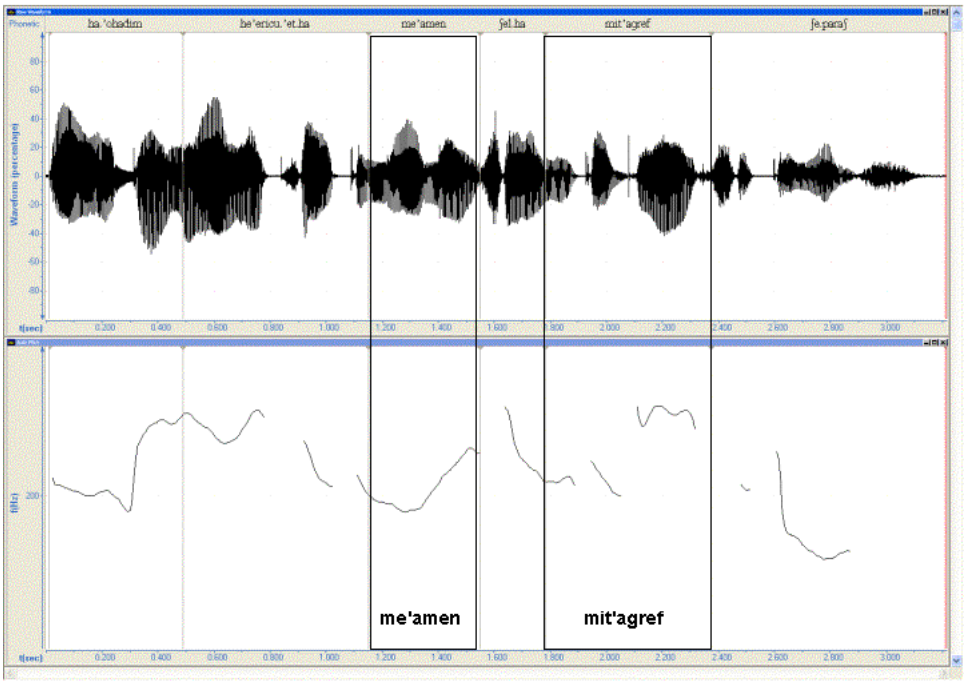


Figure 5-4: Waveform and f_0 contour for the Hebrew utterance *ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref še-paraš* 'The fans admired the coach of the wrestler who retired'. The left box indicates the N1 region, and the right box indicates the N2 region.

In measuring the f_0 contour of the N1 and N2 segments, it was important to determine whether the f_0 contour found reflected a boundary tone in addition to any f_0 shift attributable to lexical stress, as opposed to lexical stress only. As mentioned above, the boundary tone falls on the final stressed syllable, which in Hebrew is in the default case the final syllable; therefore the highest f_0 value for N1 and N2 in each utterance was obtained relative to the lowest f_0 value within that word, thus calculating 'valley-to-peak' differences. A maximum f_0 value (the 'peak') was identified within the last 100 milliseconds of the noun, as estimated in 5 x 20 millisecond bin averages. A minimum f_0 value (the 'valley') was identified within the preceding material in the same segment. Simple subtraction of the minimum f_0 value from the maximum f_0 value yielded an

'(early) valley to (late) peak' estimate of f_0 rise over the region in which a boundary tone, if present, would be expected to occur.[106] The f_0 record was manually cleaned of pitch track errors produced by the software (usually resulting in pitch halving and/or doubling); these measurements were replaced with 'corrected' estimates before pitch estimates were computed.

A duration measurement for each of N1 and N2 was estimated over a region of the waveform which included the noun and any optional pause following it, up to the beginning of the next word. Measurements reflecting pre-boundary lengthening and pausing were not taken separately, since amplitude drop-off makes it infeasible to precisely locate where a word ends and a pause begins. The analysis employed an item-based 'limiting cut-off' procedure for the duration data. This procedure identified values that were not in the range of reasonable durations, and reset them to minimize the distortion they induce in averaged data: any duration measure that exceeded twice the mean duration for that specific segment of an item (across all experimental conditions) was reset to the cut-off value. Thus any unnatural pauses and/or hesitations within a segment were sieved out so that the data consisted of a representative sample of duration measurements in Hebrew.

After thus obtaining the instrumental data from the production study, the analysis examined whether there was a correlation between the instrumental analysis and the perceived phrasing analysis, i.e., whether longer durations and higher f_0 rises were evident in regions after which native Hebrew listeners had perceived a break. For example, if a break was perceived after N2, it was expected that the instrumental

[106] In a few cases, namely for monosyllabic nouns such as *bat* 'daughter' and *sar* 'minister', peak f_0 values were extracted from three (rather than five) word-final bins.

measurements for that region would correspond to the perception of a break by exhibiting longer durations and higher pitch rises in that region. Convergence between the two modes of locating breaks provides the strongest empirical support for the differences in phrasing patterns for the different sentence types that are the focus of this investigation.

5.3 Analysis and results

As noted, two types of data analyses were performed on the elicited utterances. The first used the perceived phrasing data extracted with the ToBI- like coding system to obtain the phrasing pattern(s) associated with each experimental condition and explore differences in phrasing patterns between the different conditions. The second was an acoustic analysis that looked separately at f_0 and duration data to find correlates to the perceived phrasing analysis.

5.3.1 Perceived phrasing data

The analysis of the perceived phrasing data first looked at utterances of the 4 conditions within the high-attachment block and within the low-attachment block separately (a within-subjects analysis). This made it possible to assess the array of phrasing patterns associated with each attachment interpretation, and discover general patterns relevant specifically to the high attachment interpretation and to the low attachment interpretation. The analysis included data only from the first block that each participant encountered, i.e., for half of the participants, a high-attachment interpretation and for the other half, a low-attachment interpretation. The analysis then focused on data for the FS in its 4 conditions (created by the factorial design of RC length and attachment

height), revealing the phrasing patterns used for each condition. It examined the effect of RC length when attachment height is constant (a within-subject analysis), in order to assess whether the length of the RC has an effect on the phrasing pattern speakers choose, and the effect of attachment height on the phrasing of similar-length RCs (a between-subjects analysis) in order to examine whether a given sentence had different phrasing patterns when interpreted as attaching high versus low.[107] Finally, the analysis examined the array of phrasing patterns associated with the CS construction in its different conditions, looking for differences between the phrasing patterns of comparable conditions in FS and CS, as well as phrasing patterns similarly predominant for both constructions. (All ANOVAs for the production experiment are provided in Appendix D-1.)

A summary of the perceived phrasing data is presented in Figure 5-5 for the high-attaching utterances and in Figure 5-6 for the low-attaching utterances. The analysis focuses on the position of an intermediate phrase boundary (if any) within the complex nominal containing the RC. While the 8 experimental conditions of this study differ in their attachment interpretation (high or low attachment), construction type (FS or CS) and RC length (short or long), they do not differ with respect to their structure and segmental content prior to the complex nominal. For purposes of data analysis, the observed prosodic phrasings were grouped into three main categories based on whether they have a break after N1, after N2, or in neither position (few had a break in both positions).[108]

---

[107] Although each participant produced both variants of a sentence, one in each of the two parts of the experiment, it was decided to avoid a within-subjects comparison of the effect of attachment height on phrasing patterns, since the data could potentially be biased (see §5.2.2, and in particular footnote 93).

[108] A break after both N1 and N2 was exhibited only for CSs with a low-attaching long RC: 2.1% of the sentences in this condition exhibited a *(S V N1) (N2) (RC_1 RC_2 RC_3)* phrasing and another 2.1% of sentences in this condition had a *(SV) (N1) (N2) (RC_1 RC_2 RC_3)* phrasing (see Appendix B to this chapter).

Appendix A to this chapter shows all the specific patterns (including breaks elsewhere in the sentence) observed in the experiment that fall under each of these three categories. The most commonly occurring category of utterances exhibited the phrasing *N1 N2)(RC.* Another category of utterances exhibited the phrasing *N1)(N2 RC*. The third category had the phrasing *N1 N2 RC*, i.e., no break internal to the complex nominal.[109] The proportions of these patterns in each of the four experimental conditions are shown in Figures 5-5 and 5-6. Appendix B to Chapter 5 presents more detailed information on the incidence of these prosodic categories. In most studies, no attention is given to other positions except for the critical region; in this study, all possible break positions were considered, as can be seen in Appendix B to this chapter, since they might turn out to be relevant. As a matter of fact, other break positions observed in the sentence (prior to the critical region, and within the RC) do not seem to make a difference; therefore we grouped the perceived phrasing into the three categories below. (See Appendix C to this chapter for an analysis of perceived breaks throughout the sentence.)

This phrasing pattern was not found in any of the other conditions; it constitutes 1.05% of all utterances produced in the experiment.

[109] I follow Selkirk (2000) in using parentheses to indicate a major phonological phrase boundary (as opposed to square brackets, which are used to indicate syntactic structure).

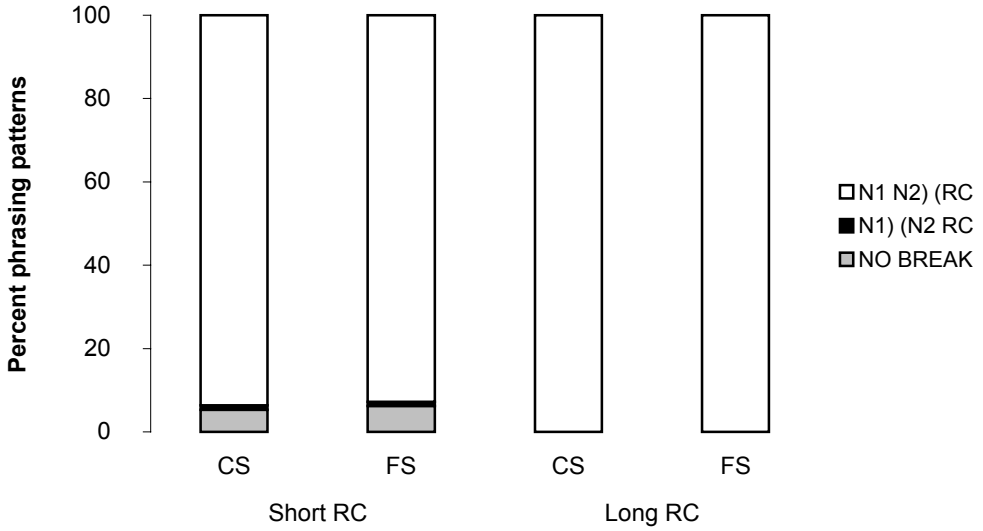


Figure 5-5: Percentages of perceived breaks for high attachment utterances, categorized by prosodic break position, as a function of construction type and RC length.

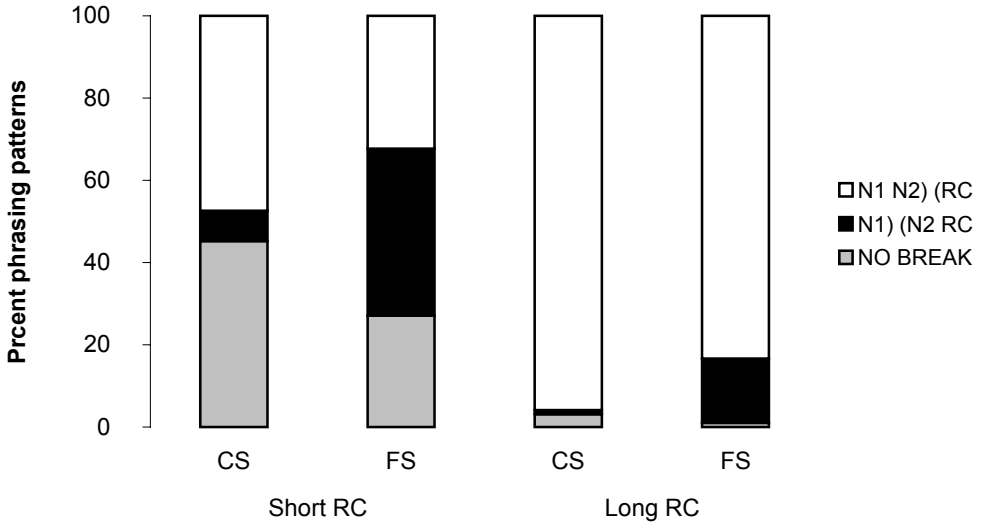


Figure 5-6: Percentage of perceived breaks for low attachment utterances, categorized by prosodic break position, as a function of construction type and RC length.

The data for the forced high attachment utterances reveal that participants used almost exclusively the same prosodic phrasing category (with a break between N2 and the RC), regardless of construction type and RC length. A within-subjects analysis of the high attachment data shows that breaks between N2 and RC predominated in all 4 conditions: FS-short RC (92.8%), FS-long RC (100%), CS-short RC (93.7%) and CS-long RC (100%). There was no main effect of construction type and no interaction (F<1.0). The item-based analysis did reveal an effect of RC length (a reduced likelihood of a break after N2 when the sentence contained a short RC, corresponding to increased likelihood of incorporating the RC into the prosodic phrase containing the complex nominal), but it was only a trend in the subject-based analysis (F_1 (1,11)=4.12, .05<p<.1; F_2 (1,7)=10.72, p<.05), suggesting that this phrasing strategy was not used by all participants consistently. Within the specific prosodic patterns that fall under the category *N1 N2)(RC*, the phrasing pattern that was used overwhelmingly in all 4 high attachment conditions was *(SV) (N1 N2) (RC)*, in which the subject and verb are grouped into one prosodic unit, the complex nominal into another prosodic unit, and the RC (short or long) into a prosodic unit on its own. This was the pattern for 89.8% of all high attachment utterances.[110] Appendix B to this chapter shows other phrasing patterns that are subsumed under this category, and their relative frequency.

In the low-attaching condition, by contrast, the results show variation in the phrasing patterns for the 4 experimental conditions. A general distinction can be made

[110] The *(SV) (N1 N2) (RC)* pattern was used 86.5% with FS-short RC, 94.7% with FS-long RC, 86.3% with CS-short RC and 91.7% with CS-long RC. The alternative phrasing pattern *(S) (V N1 N2) (RC)* was used an average of 5.3% over all conditions. Two other phrasing patterns were used with a pre-RC break: one had no additional break prior to the *N2)(RC* break (1% of CS-long sentences) and the other used the *(S V) (N1 N2) (RC)* pattern, but had an additional break internal to the RC (2.1% of FS-long RC sentences, 1% of CS-long RC sentences; see Appendix B to Chapter 5 for compete details).

between the phrasing patterns exhibited for long RCs and those exhibited for short RCs: the *N1 N2)(RC* category dominated the long RC conditions (for both CS and FS sentences), as it did for the high-attaching utterances. This was expected (on the assumption that breaks within the RC tend to be avoided)[111] since a long RC of 3 prosodic words is heavy enough to constitute its own prosodic unit. [112] [113] The data further show that the category *N1 N2)(RC* was less frequent in the short RC cases, compared with the long RC cases. A comparison of this prosodic category across the 4 conditions shows a main effect for RC length: there were more breaks after N2 for sentences with a long RC than those with a short RC (F_1 (1,11)=29.89, p<.001; F_2 (1,7)=190.22, p<.001).[114] Low-attaching short RC sentences exhibited the use of several alternative phrasing patterns, without a clear preference for one phrasing pattern; although the array of phrasing patterns found for FS and CS with short RC is different, this can be explained as following from their phonological properties, as will be discussed below.

---

[111] Possibly due to favoring placement of prosodic breaks at clause boundary positions over other [XP boundary positions (such as a [PP boundary).

[112] RC-internal breaks always occurred following the first prosodic word within the RC (the subordinate verb); In the low attachment cases, several phrasings exhibited an RC-internal break, but they were all subsumed under the *N1)(N2 RC* category (3% of FS sentences and 1% of CS sentences). In the high-attaching cases, they occurred in 2.1% of the FS sentences and in 1% of the CS sentences; the pattern used was (SV) (N1 N2) (RC) (RC RC), in which there is a break both before the RC and within the RC. Overall, RC internal breaks were found in only 2.3% of the sentences produced in this study. See Appendix B to this chapter for more details.

[113] A long RC is likely to induce a prosodic break before it for phonological length reasons. Watson & Gibson (2004) have proposed that as the size of an upcoming constituent increases, the likelihood of boundary placement before it increases also, for processing reasons (memory requirements for planning the next linguistic unit). D'Imperio et al. (2005) argue against this explanation, claiming that there might be some processing effects, but in addition there are language specific phonological requirements, as they showed for Catalan.

[114] An effect of construction type was also observed for the *N1 N2)(RC* category (F_1 (1,11)=10.16, p<.01; F_2 (1,7)=36.83, p<.001), with more breaks after N2 produced for CS than for FS. There was no interaction between the two main effects. This will be discussed in relation to the CS data below.

Let us consider the results for the FS first. A summary of the perceived phrasing data is presented in Figure 5-7, which recasts the data for the 4 conditions of the FS created by the RC length and attachment height factors, previously presented in Figures 5-5 and 5-6.

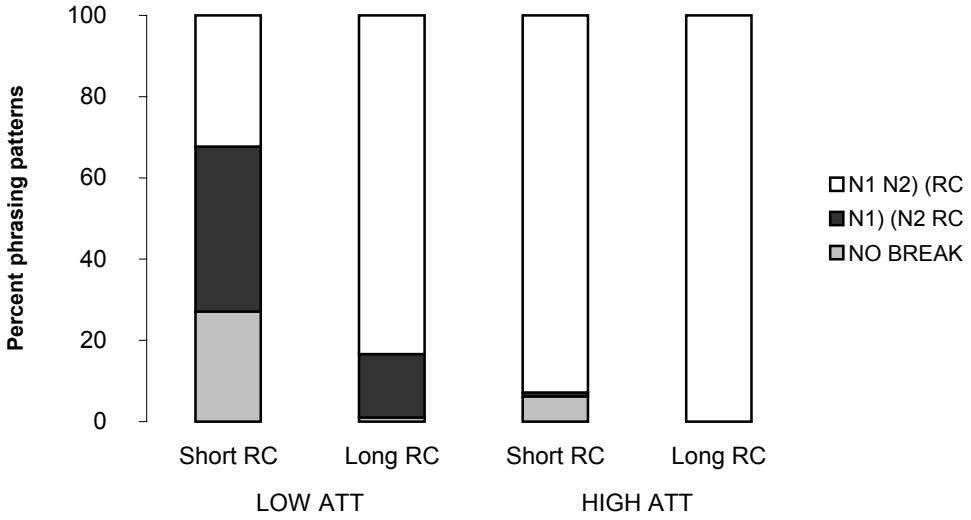


Figure 5-7: Percentage of perceived phrasing for FS utterances, categorized by prosodic break position, as a function of attachment height, and RC length.

A comparison of the *N1 N2)(RC* category for the 4 conditions of the FS revealed an effect of attachment height (F_1 (1,11)=22.96, p<.001; F_2 (1,7)=108.54, p<.001), with this phrasing pattern occurring more often with high-attaching utterances than with low-attaching utterances. There was also an effect of RC length (F_1 (1,11)=32.23, p<.001; F_2 (1,7)=841.00, p<.001): the *N1 N2)(RC* phrasing pattern was predominantly used with long RCs, and to a significantly lesser extent with the short RCs. There was an interaction of the two main effects (F_1 (1,11)=37.71, p<.001; F_2 (1,7)=84.00, p<.001). A sub-analysis looking separately at high and low attachment interpretation therefore followed.

As noted above, high attachment utterances overwhelmingly had an intermediate phrase boundary immediately following N2, regardless of RC length; there was no statistical difference between short and long RCs in the occurrence of the *N1 N2)(RC* pattern in the subject-based analysis, although the difference was significant in the item-based analysis (F_1 (1,11)=2.57, p>.1; F_2 (1,7)=5.65, p<.05). The somewhat lessened use of this pattern in the high-attaching short RC case was complemented by use of the alternative pattern with no break within the N1 N2 RC region. The predominance of the *N1 N2)(RC* pattern was exhibited to almost the same extent in the low-attaching long RC condition. However, although there was no statistical difference between this condition and its equivalent high-attaching condition in the subject-based analysis (p>.1), the difference was significant in the item-based analysis (F_2 (1,7)=18.02, p<.005). Utterances with a low-attaching short RC, in contrast to the three other conditions, exhibited an array of prosodic patterns, as discussed below.

Focusing now on the low attachment data, FS utterances with a long RC predominantly used the *N1 N2)(RC* pattern, but other patterns were found as well: 15.6% of FS utterances had a break after N1 rather than between N2 and RC, and 1% had no break at all in the complex NP with RC.[115] The *N1 N2)(RC* pattern was used less frequently with a short RC than with a long RC (F_1 (1,11)=38.48, p<.001; F_2 (1,7)=248.90, p<.001); in its place, the *N1)(N2 RC* pattern was observed in greater frequency in the short RC than in the long RC condition (F_1 (1,11)=10.56, p<.01; F_2

[115] Subjects who used the *N1)(N2 RC* pattern may have chosen this phrasing category in order to convey unambiguously the low attachment interpretation. However, the length of the RC made it unlikely for the material following the N1 break to remain a single prosodic unit. Indeed, only 8.4% exhibited such a pattern. The remaining were divided between cases in which there was an additional break after N2, creating an *N1)(N2)(RC* pattern (4.2%), and cases in which an additional break was used inside the RC (3%).

(1,7)=42.00, p<.001), and there was also a substantial percentage of utterances with no break after either N1 or N2. Thus, FS sentences with a low-attaching short RC showed a mixed pattern of results: speakers produced utterances falling into all three of the phrasing categories, with no statistical preference among them, $\chi^2(2)=2.688$, p>.10.[116] This range of phrasing patterns was markedly different not only when compared with the low-attaching long RC, but also when compared to the high-attaching short RC (F_1 (1,11)=36.71, p<.001; F_2 (1,7)=117.71, p<.001); while the *N1 N2)(RC* pattern was predominant in the high-attaching short RC condition, it was only one of several equally plausible patterns exhibited in the low attachment short RC condition.

We now turn to the CS data. Figure 5-8 presents the summary phrasing data for the 4 conditions tested with the CS in the production study, recast from the original phrasing data presented in Figures 5-5 and 5-6.

[116] In each experimental condition 12 participants uttered 8 items, for a total of 96 utterances per condition. In the High-attaching long RC condition, there was a missing data point, thus the percentage was calculated with a total of 95 utterances.

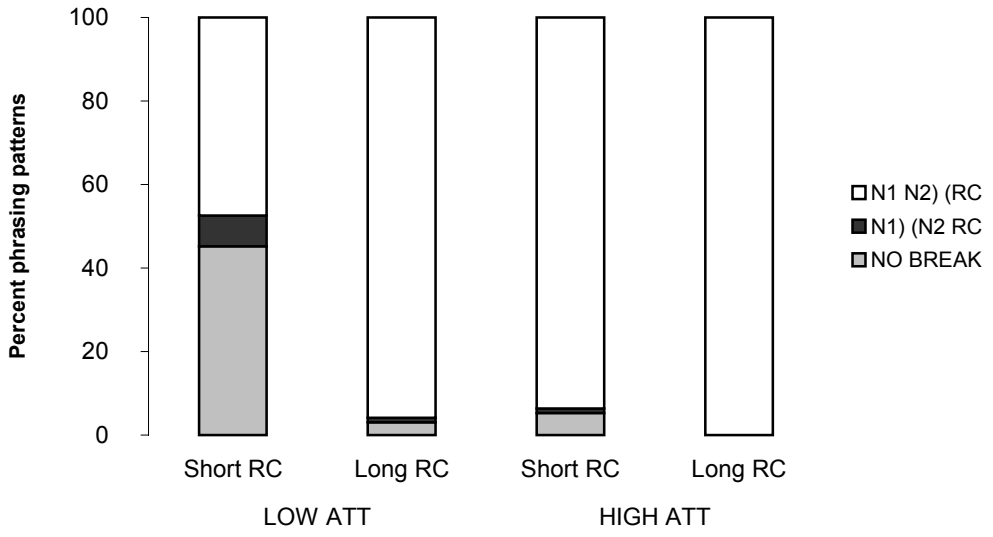


Figure 5-8: Percentage of perceived phrasing for CS utterances, categorized by prosodic break position, as a function of attachment height, and RC length.

As with the comparable FS conditions, CS utterances with high-attachment meaning predominantly used the *N1 N2)(RC* pattern, regardless of RC length (F_1 (1,11)=3.67, .05<p<.1; F_2 (1,7)=4.20, .05<p<.1). The low-attaching conditions showed variation of phrasing patterns between short and long RC: For long RCs attaching low, as in the FS sentences, the *N1 N2)(RC* category dominated. However, the exceptions were fewer for the CS than for the FS: 1% with a break after N1 (cf. 15.6% in FS) and 3.1% with no break (cf. 1% in FS). The fewer breaks after N1 in the case of the CS provide support for the standard claim that the CS constitutes a single prosodic word. Turning now to the case of a short RC attaching low to a CS, there was a clear preference for no break, or a break between N2 and RC, rather than a break between the nouns (χ^2(2)=28.884, p<.001). There was no preference between the no break phrasing and the phrasing with a break between N2 and RC (χ^2(2)=0.045, p>.10). There were breaks

between the two nouns in just 7.4% of the utterances in the short RC low attachment condition. These were produced by 2 of the 12 subjects whose utterances had a low-attachment meaning; it is possible that they used this unnatural prosody as a disambiguation strategy, in order to convey unambiguously a low-attachment interpretation.[117] Despite those occasional exceptions, these data offer strong support for the standard claim that the CS is a single phonological unit and no break is permitted within it.

Comparing FS and CS utterances of low-attaching short RCs shows that the *N1 N2)(RC* pattern was used more frequently with CS than with FS (F_1 (1,11)=7.59, p<.05; F_2 (1,7)=22.87, p<.005). This is to be expected, given the different characteristics of the FS and the CS: since a break between the two nouns is not permissible in the CS, it is reasonable to expect participants to resort to other phrasing patterns, among them the category in which a break occurs after N2. The data clearly supports the prediction that the range of possible phrasing patterns available for CS and FS is not the same: the 'no break' pattern is available for the FS, and is used as frequently as the other alternative phrasings (in the low-attaching short RC condition, the only condition in which great variation in phrasing patterns is exhibited), while it is rarely used in the CS conditions. It is possible that the variation in available phrasing patterns for the two constructions has an affect on the attachment interpretation of these sentences in silent reading, as will be discussed in Chapter 7.

To summarize: The analysis of the perceived phrasing data reveals variation in phrasing patterns only for sentences with forced low attachment interpretation; high

[117] In the high attachment conditions, there was 1 utterance, in the CS-short RC condition, that had a break between N1 and N2; however, in this case the break was characterized as a rise accompanied by a long pause, which was perceived as a hesitation rather than an intermediate prosodic break.

attachment interpretation was reflected overwhelmingly by the use of the *N1 N2)(RC* phrasing category. The data showed a length effect only for sentences interpreted as attaching low, with the use of the *N1 N2)(RC* phrasing predominantly with long RC, and a mixture of different phrasings for cases with short RC. Finally, the data showed that the phonological properties of the CS indeed restrict the range of phrasing patterns associated with it. While the FS-short RC condition freely used all three phrasing categories when interpreted low, the CS-short RC utterances omitted almost entirely the phrasing category in which a prosodic boundary would divide the nouns of the CS, confirming the assumption that the grammar requires it to be a single prosodic unit. The short RC cases also reveal that within the observed range of phrasing patterns permissible for each construction, there is no clear preference for one phrasing pattern over another.

5.3.2 Acoustic data

Results for the f_0 measure are presented in Figures 5-9 and 5-10. Figure 5-9 presents mean values for N1 "rise" and Figure 5-10 presents mean values for N2 "rise". Results for the duration measure are presented in figures 5-11 and 5-12. Figure 5-11 presents N1 duration means and Figure 5-12 presents N2 duration means. (All ANOVAs for the production experiment are provided in Appendix D-1.)

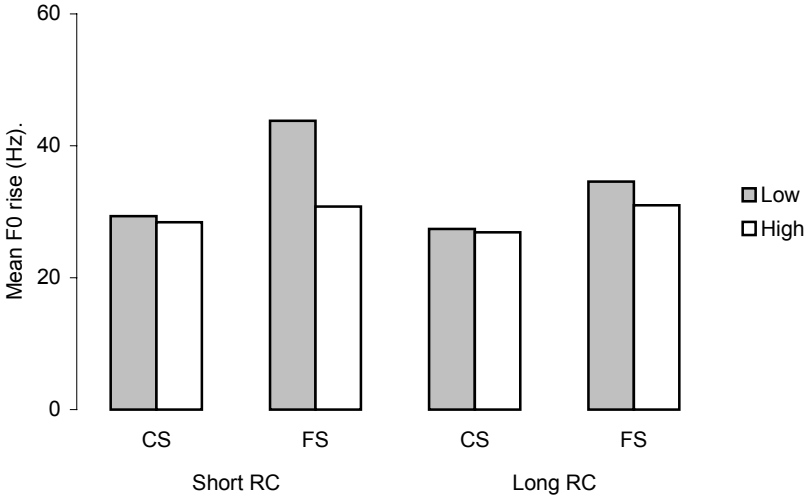


Figure 5-9: Valley-to-peak f_0 rise for short and long RC, as a function of construction type, RC length and attachment height, for N1

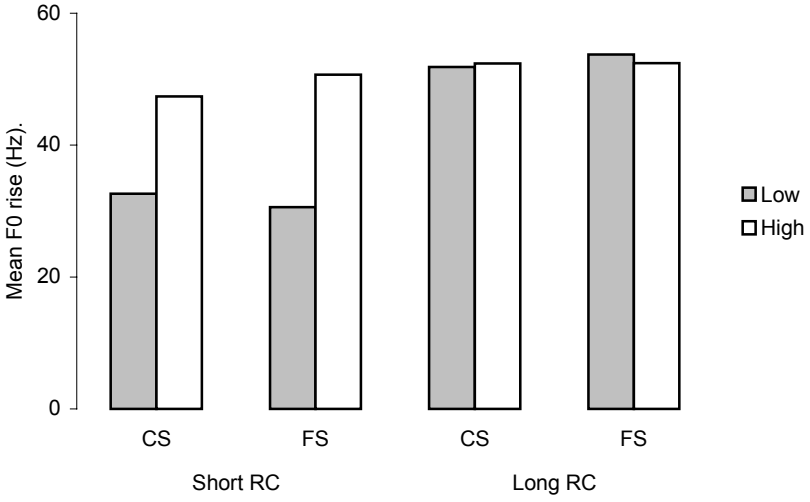


Figure 5-10: Valley-to-peak f_0 rise for short and long RC, as a function of construction type, RC length and attachment height, for N2

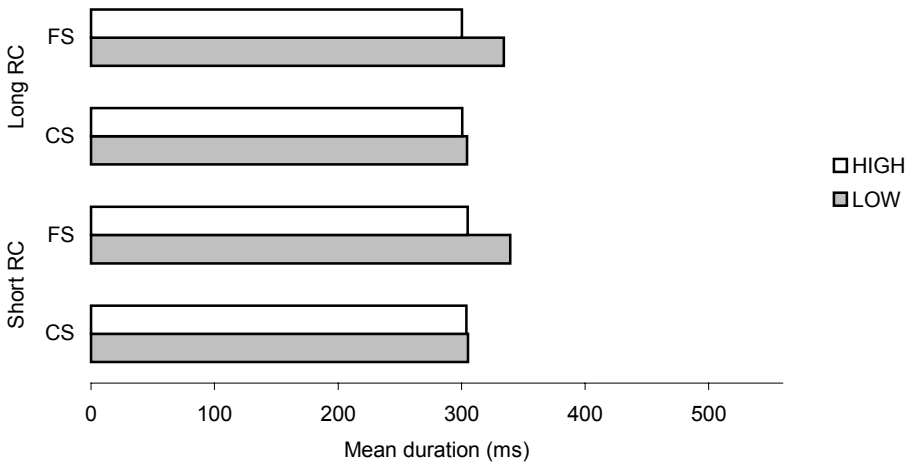


Figure 5-11: Duration of the N1 region, as a function of construction type, RC length and attachment height.

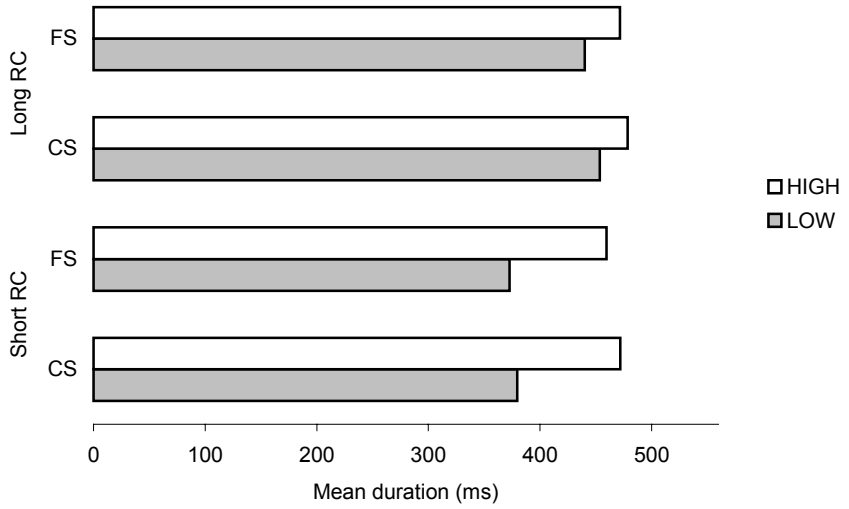


Figure 5-12: Duration of the N2 region, as a function of construction type, RC length and attachment height.

Looking first at the high attachment data in general, in comparison to the low attachment conditions, shows that high attachment yielded higher f_0 rises and longer durations at the N2 segment, suggesting that a phrase boundary occurred after N2, as was inferred also on the basis of the perceived phrasing data. This was complemented by lack of a rise or lengthening at N1 in all high attachment conditions, suggesting the absence of a prosodic break following N1. A comparison of the f_0 data for high attachment across conditions shows that for both N1 and N2, there was hardly any variation, i.e., there was no RC length effect, no construction type effect and no interaction (all p's >.1). The average overall rise at N2 was 51Hz and the average overall rise at N1was 29Hz. Similarly, the duration data for high attachment show no effects of the conditions and no interaction: For N2 Duration, there was no effect of RC length (F's<1.0), no effect of construction type (F's<1.0), and no interaction (F's<1.0). For N1 duration, as well, there was no effect of RC length (p's>.20), no effect of construction type (p's>.20) and no interaction (F's<1.0). The overall duration of N1 was 303 ms and the overall duration of N2 was 471 ms.[118] Both f_0 and duration data thus support the perceived phrasing data and together show that in producing the high attachment interpretation of the target construction, Hebrew speakers overwhelmingly used a prosodic phrasing with a break following N2 and preceding the RC, whether the RC is short or long, and whether it attaches to an FS or a CS.

FS: f_0 data. Analysis of the f_0 data for N1 in the FS cases revealed an average rise over N1 that was 13 Hz greater for low-attaching short RC sentences than for their high-

[118] The difference in overall f_0 rise and length between N1 and N2 was not surprising, due to lexical variation. See §2.5.2 for more details.

attaching equivalents (F_1 (1,11)=3.98, .05<p<.1; F_2 (1,7)=22.01, p<.005). No such difference in the f_0 rise for N1 was seen in the low- versus high-attaching sentence types with long RC (p's>.20); this provides a useful control for possible speaker group differences, since it showed that the effect is not a result of a between-subjects comparison (comparisons between sentences with high- and low-attached RC are between speakers whereas long versus short RC is a within-speakers comparison). It is noteworthy that the average f_0 rise in N1 of all long RC sentences (33 Hz) is very like that of high-attaching short RC sentences (30 Hz), suggesting that in these conditions, there is no boundary tone suggestive of the end of a prosodic phrase. Together, these data suggest the presence of a prosodic break after N1 in the low-attaching short RC condition, and the absence of such a break in the three other conditions. For N2 the f_0 data are complementary to those for N1: f_0 rise was 20 Hz greater for a high-attaching short RC than for its low-attaching equivalent, F_1 (1,11)=5.72, p<.05; F_2 (1,7)=16.12, p<.01. And once again, the data for long RCs (i.e., within-subject control) show no such difference (p's>.50), and their average f_0 rise (53 Hz) accords with that for high-attaching short RC sentences (51 Hz). It can be concluded that a prosodic break occurred after N2 for both long RC conditions and in the high-attaching short RC condition, but not in the low-attaching short RC condition.

FS: duration data. A contrast between a low- and high-attaching short RC can be observed in the duration measures for N1: N1 was significantly greater for the low-attaching short RC (F_1 (1,11)=5.18, p<.05; F_2 (1,7)=38.18, p<.001), which is compatible with the occurrence of a prosodic break after N1 for a low-attaching short RC, but not for

a high-attaching short RC. On the basis of the perceived phrasing data and the f_0 data, it could be expected that a break after N1 would show up in the duration measure *only* for the low-attaching short RC. However, this was not the case. Contrary to expectations, N1 duration was also longer for the low-attaching than the high-attaching *long* RC (F_1 (1,11)=3.50, .05<p<.1; F_2 (1,7)=8.66, p<.05). In fact, the low-attaching long RC did not significantly differ in N1 duration from the low-attaching short RC (p's>.1). This is a rare discrepancy between the duration data and the perceived phrasing data (and also between the duration and f_0 data); at present it is not clear why this is so. For the N2 duration data, a comparison between the low-attaching short RC condition and the three other conditions does conform to the other measures. That is, N2 was longer for a high-attaching RC (both short and long) and for a low-attaching long RC, compared to the low-attaching short RC. This suggests that a prosodic break after N2 occurred in all conditions except for the low-attaching short RC. Comparing high- and low-attaching short RC shows that the duration of N2 for a high-attaching short RC was significantly greater than that for a low-attaching short RC (F_1 (1,11)= 10.91, p<.01; F_2 (1,7)=77.01, p<.001). This suggests that, as found in the f_0 and perceived phrasing data, a prosodic break after N2 is more probable for a high-attaching short RC than for a low-attaching short RC.

CS: f_0 data. For the CS cases, N1 rise was similar for both short and long RC, in both high and low interpretations, with an average rise of 28 Hz (CS-short RC-low 29.5Hz; CS-long RC-low CS 27.4Hz; CS-short RC-high 28.4Hz; CS-long RC-high 26.9 Hz) (p's>.1, except for an RC length effect in the item-based analysis, F_2 (1,7)=6.43, p<.05).

Note that this rise is quite similar in magnitude to that found for the N1 rise of FS low-attaching long RC (33 Hz), and high-attaching short RCs (30 Hz), where it was concluded that a phrasal break did not occur (see above). Thus, it is possible to conclude from this comparison that no phrasal rise occurred after N1 in the CS conditions. The absence of a rise indicating the occurrence of a phrasal break following N1 of the CS was expected, due to the status of the CS as a single phonological word. N2 in the CS conditions shows variation, in accord with findings from the phrasing analysis: N2 rises (average 50.5 Hz) were higher for high-attaching RCs and for a low-attaching long RC. The N2 rise was higher for a long RC attaching low compared with a short RC attaching low (F_1 (1,11)=8.85, p<.05; F_2 (1,7)=27.42, p<.005), while there was no statistical difference between the low-and high attaching long RC conditions (p's=1.0). This N2 rise of 50.5 Hz on average is quite similar to the N2 rise observed in comparable conditions of the FS, for which it was concluded that it signifies a prosodic break.

CS: duration data. The duration data of N1 in the CS conditions, similarly to the f_0 data of N1 for the CS, show no variation: the average duration of N1 was 304 ms (cf. durations of the FS in those conditions in which no lengthening was observed: 304ms and 300ms for short and long RC attaching high, respectively), suggesting there was no lengthening of N1 signifying a prosodic break in the CS cases, as expected. There was no effect of height of attachment, no RC length effect and no interaction (all p's >.1). N2 duration data show variation similar to that of the N2 rise data: N2 duration was shorter for a short RC attaching low than for a short RC attaching high (F_1 (1,11)=5.55, p<.05; F_2 (1,7)=67.80, p<.001), while the short and long RCs attaching high did not differ

significantly (p's>.1). In other words, N2 lengthening was observed where it was expected (average duration 468 ms; cf. average of 457 ms in the equivalent FS cases). It did not occur where it was not expected, in the short RC attaching low (average N2 duration 380 ms; cf. 373 ms in the comparable FS condition).

Instrumental data support phrasing data. The instrumental data provide strong confirmation of the perceived phrasing analysis: in all but the low-attaching short RC conditions, the f_0 rises and greater durations at N2 correspond very closely to the perceived phrasing. For the FS, the two acoustic measures (f_0 and duration) indicate a greater probability (or strength) of prosodic breaks after N1 for the low-attaching short RC, but not for the other three conditions. In line with the results of the perceived phrasing analysis, N1 had greater f_0 and duration measures for a low-attaching short RC, compared with a high-attaching short RC. f_0 rise measures for N1 were also greater for a low-attaching short RC than for low- and high-attaching long RCs although this was not fully supported by the duration data. A complementary pattern was found for N2: there was an f_0 rise and lengthening in all but the low-attaching short RC. This suggests, in accordance with the perceived phrasing findings, that the phrasing pattern *N1 N2)(RC* was predominant in all conditions except the low-attaching short RC. The acoustic data for the FS are also consistent with the finding of a more mixed pattern in the low-attaching short RC condition: in the phrasing data, roughly a third of the utterances exhibited the phrasing *N1)(N2 RC*, and this condition was the only one where acoustic measures indicated incidence of a break after N1 and weaker signs of a break after N2.

For the CS conditions, the acoustic data show a lack of lengthening and f_0 rise on N1, corresponding to the finding in the perceived phrasing data that this pattern was rarely used in this construction, and confirming the assumption that the CS is a single phonological word with no break permissible between its two nouns. The N2 data for the CS show shorter duration and smaller rises on the short RC attaching low, suggesting that no break occurs after N2 in this condition, which supports the findings of the perceived phrasing data. The remaining conditions of the CS exhibited longer durations and higher rises, suggesting that a break occurred after N2 in these conditions, as was found in the perceived phrasing analysis. Thus the analysis of the acoustic measures for both the CS and FS conditions found clear correlates of the results obtained from the perceived phrasing data: the predominant *N2)(RC* breaks for high (short and long RC), the RC length effect for low attachment only, and the markedly fewer breaks between N1 and N2 for CS in the low attachment condition.

Lexical vs. Phrasal stress. The f_0 data contribute further to our understanding of the difference between lexical and phrasal stress in Hebrew. Recall that Hebrew places lexical stress by default on the word-final syllable. The final syllable is also the location where phrasal stress is marked. The f_0 data suggest that listeners indeed abstract away from the lexically intrinsic rise in identifying prosodic phrase pitch rises (boundary tones). Informal estimations (derived from linear regressions) comparing perceived phrasing data and f_0 data showed that for cases in which a phrasal break was perceived, pitch increased by some 35 Hz above the level observed when no prosodic break was perceived. Based on estimates of intrinsic vowel rise (see below), a combined analysis of

the f_0 rise for cases where a prosodic break was perceived in the phrasing data revealed that the average f_0 rise for N1 was approximately 36 Hz with a 100% increase in likelihood of (perceived) break. The baseline (zero intercept) is 30 Hz, a value which reflects lexical stress only. For N2, the average f_0 rise was 34 Hz with a 100% increase in likelihood of (perceived) break. The baseline (zero intercept) is 20 Hz, which reflects lexical stress only. The different intercepts for N1 and N2 have to do with change in the realization of f_0 of different lexical items, which is modulated by the vowel qualities over which valley and peak measurements are instantiated.

Most et al. (2000) studied articulatory characteristics of Hebrew vowels. Their findings of the means f_0 (Hz) of the 5 Hebrew vowels produced by women were used to estimate the f_0 rise due to intrinsic vowel height in N1 and N2 of the materials for this study. A preliminary analysis indeed shows that N1 words had higher intrinsic vowel rise than the N2 words, which may be the cause for the different intercepts seen in our informal comparison. Although there is no data on the actual vowel qualities of the speakers of the production experiment reported here, there is every reason to expect that they would be similar to those of the sample collected in Most et al. (2000). In sum, the acoustic analysis suggests that lexical and phrasal stress in Hebrew are realized differently in terms of acoustic measures, with phrasal stress realized as an average rise of 35 Hz over the intrinsic lexical stress. This finding also contributes to an assessment of the reliability of the perceived phrasing judgments, since it suggests that indeed Hebrew speakers can abstract away from intrinsic lexical stress in identifying prosodic phrasing of spontaneous utterances.

5.4 Summary and conclusions

The analyses of data from the production study revealed a variety of prosodic phrasing patterns used in expressing high and low attachment meanings across the experimental conditions. Prosodic phrasing varied as a function of attachment height (forced high/low RC-attachment). An FS with a short RC exhibited 92.8% breaks before the RC for high attachment, but only 32.3% for low attachment. This is predicted by Selkirk's *Align*$_{\boldsymbol{R}}$ *XP* constraint, which would require a break at the right brackets of the maximal projections]DP]PP that precede a high-attaching RC (see §5.1 for full syntactic structure). The fact that identical word-strings, disambiguated for attachment height only by the preamble question, showed strongly contrasting prosody provides clear validation for the Post-to-Times elicitation protocol, which has not previously been used with high attachment in any language.

As has been found in Croatian (Lovrić, 2003), an increased number of prosodic breaks were found at the left edge of the relative clause when it was long, than when it was short, in line with the hypothesis that a pre-RC break is often the default phrasing for a long RC (Fodor, 2002). However, in Croatian the effect of RC length did not interact with an effect of attachment height: in both high and low attachment utterances, there were more breaks before a long RC than before a short RC. In Hebrew this happened in the low attachment productions only. For high attachment, pre-RC breaks predominated in all conditions, and equally so. The data suggest that differences in prosodic phrasings of the experimental conditions are driven by a ranking of constraints specific to Hebrew. Shaked (2007) considers the ranking of constraints on prosodic phrasing in Hebrew using an analysis of part of the production data reported in this dissertation, specifically,

utterances of the FS in the 4 experimental conditions created by attachment height (high/low) and RC length (short or long RC). It was proposed there that in Hebrew, the alignment constraint is ranked higher than the length constraints. The perceived phrasing data and the acoustic data together show that the alignment constraint applies without regard to the phonological length constraints. For the high-attachment conditions, the alignment constraint overwhelmingly triggered a break after N2, immediately preceding the RC, a position in which there is a discontinuity in the syntactic tree. A prosodic break between N2 and RC, leading to the RC constituting a single prosodic unit of 1 prosodic word in the short RC case and 3 prosodic words in the long RC case, was observed even though for a short RC this results in a violation of the optimal length constraint *BinMin* (requiring a minimum of 2 prosodic words in a prosodic phrase), and for a long RC it violates the *BinMax* constraint (requiring a prosodic phrase to consist optimally of a maximum of 2 prosodic words).

In the low attachment cases there are no syntactic maximal projection right edges after N2, and therefore the alignment constraint is inapplicable, so there the interplay of the length constraints can be observed. The results suggest that the length constraints *BinMin* and *BinMax* do play a role in prosodic phrasing of Hebrew utterances, but neither strongly dominates the other: a short RC attaching low was evidenced to the same extent with violations of *BinMin* (e.g., the *N1 N2)(RC* phrasing) and *BinMax* (e.g., the phrasing of N1 N2 and RC into a single prosodic phrase), with no clear dominance for one over the other. It is thus concluded that Hebrew speakers have a preference for phrases which are optimally of 2 prosodic words in length; see further discussion in Shaked (2007). This is particularly interesting when comparing Hebrew to other languages in which *BinMin*

and *BinMax* are less balanced. For example, in Standard Arabic, *BinMin* has been shown to be less dominant than *BinMax*, as Abdelghany's (in preparation) study of the RC construction in Standard Arabic reveals. In the Standard Arabic data on the RC construction, the minimum size requirement for a Major Phrase is 4 prosodic words and it is proposed that *BinMin* in Standard Arabic is a constraint requiring a major phrase to consist of at least four prosodic word (*BinMin* is interpreted in Abdelghany's study as a categorical/gradient constraint; each violation of the constraint will accrue the respective number of violations). *BinMin* is presented as a weak constraint in Standard Arabic; Abdelghany proposes instead a **P-Phrase* constraint, which prefers fewer and larger p-phrases. Hellmuth (2004) also shows, on a basis of a variety of constructions in Colloquial Egyptian Arabic, a preference for long phrases, and subsequently suggests the ranking of *BinMax* above *BinMin*.

For the RC construction under investigation, it is interesting to observe similarities and differences in the productions of the FS and CS. For a long RC, both constructions show a clear preference for a break between N2 and the RC, regardless of attachment height. Although utterances were produced for contextually disambiguated versions of the ambiguous RC construction, it is possible to conclude that an ambiguous long-RC sentence would also be assigned this prosodic contour, since it dominated all long-RC utterances regardless of attachment. This suggests that *N1 N2)(RC* is the default prosodic pattern for the RC construction when RC is long. The large amount of pre-RC breaks found in the production study would thus suggest that Hebrew would be a 'high attaching' language for Long RC, as argued for French (Quinne et al., 2000). This prediction will be addressed in Chapter 7, §7.5.

For a short RC, there was no consistent phrasing pattern observed throughout all conditions tested in this experiment. The high-attaching short RC predominantly used the *N1 N2)(RC* pattern in both FS and CS, while the low-attaching short RC exhibited a variety of patterns. Among the low-attaching short RC, a difference between FS and CS was found, with 40.6% breaks between N1 and N2 for short-RC FS but only 7.4% for short-RC CS. This difference can be ascribed to the phonological word constraint on CS, which disfavors a break between the two nouns. In general, since short utterances of the RC construction, both with FS and CS, did not exhibit a consistent phrasing pattern, it can be concluded that there is no default prosodic structure for such sentences, and the choice between the available prosodic phrasing patterns for each construction with this structure are either random or are due to other factors not controlled in this study.

The production study revealed the variety of phrasing patterns used to express disambiguated versions of the RC sentences. High attachment interpretation was shown to be associated with a break preceding the RC, regardless of the RC length; a long RC was shown to be associated with a break preceding the RC, regardless of attachment. We now turn to investigate the perception of fully ambiguous sentences; this will be examined in a listening study where sentences are presented auditorily with varying prosodic structures (Chapter 6) and in a silent reading experiment where no prosody is provided in the input (Chapter 7).

Appendix A to Chapter 5: Variety of phrasing patterns observed in the production study and their assignment to three phrasing categories.

Phrasing patterns perceived in participants' productions are shown here with specific patterns that fall into each of the three phrasing categories established for analysis. Individual patterns within a category are listed in decreasing order of frequency. See Appendix B to this chapter for additional details. RC_1 designates the short RC. RC_2 and RC_3 designate the two additional minor phrases in the long RC. Patterns marked * have a prosodic break both after N1 and after N2. Intuitively, the break after N1 rules out the high-attachment interpretation of the sentence. It is also the first break in the critical region, suggesting that this was the initial prosodic pattern intended by the speaker. For these reasons, phrasing pattern was categorized under the *N1)(N2 RC* rubric.

N1 N2 RC	**N1) (N2 RC**	**N1 N2) (RC**
(S V) (N1 N2 RC_1)	(S V) (N1) (N2 RC_1)	(S V) (N1 N2) (RC_1 RC_2 RC_3)
(S) (V N1 N2 RC_1)	(S V N1) (N2 RC_1)	(S V) (N1 N2) (RC_1)
(S) (V) (N1 N2 RC_1)	(S) (V N1) (N2 RC_1)	(S) (V N1 N2) (RC_1)
(S V) (N1 N2 RC_1 RC_2 RC_3)	(S V N1) (N2 RC_1 RC_2 RC_3)	(S) (V N1 N2) (RC_1 RC_2 RC_3)
	(S V N1) (N2) (RC_1 RC_2 RC_3)*	(S V) (N1 N2) (RC_1) (RC_2 RC_3)
	(S) (V N1) (N2 RC_1 RC_2 RC_3)	(S V N1 N2) (RC_1 RC_2 RC_3)
	(S V) (N1) (N2) (RC_1 RC_2 RC_3)*	
	(S V N1) (N2 RC_1) (RC_2 RC_3)	
	(S) (V N1) (N2 RC_1) (RC_2 RC_3)	
	(S V) (N1) (N2 RC_1) (RC_2 RC_3)	

Appendix B to Chapter 5: Incidence of phrasing categories established for the analysis of the production study.

Observed phrasing patterns for each of the four conditions (RC length x attachment height) are presented here for CS and for FS, divided into the three main phrasing categories employed in the phrasing analysis (see Appendix A to this chapter), with the percentage of occurrence for each phrasing pattern.

High-attaching short RC

FS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1)	4.2%
	(S) (V N1 N2 RC_1)	1%
	(S) (V) (N1 N2 RC_1)	1%
N1) (N2 RC	(S V N1) (N2 RC_1)	1%
N1 N2) (RC	(S V) (N1 N2) (RC_1)	86.5%
	(S) (V N1 N2) (RC_1)	6.3%

CS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1)	5.3%
N1) (N2 RC	(S V) (N1) (N2 RC_1)	1.1%
N1 N2) (RC	(S V) (N1 N2) (RC_1)	86.3%
	(S) (V N1 N2) (RC_1)	7.4%

Low-attaching short RC

FS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1)	21.9%
	(S) (V N1 N2 RC_1)	5.2%
N1) (N2 RC	(S V) (N1) (N2 RC_1)	20.8%
	(S V N1) (N2 RC_1)	14.6%
	(S) (V N1) (N2 RC_1)	5.2%
N1 N2) (RC	(S V) (N1 N2) (RC_1)	29.2%
	(S) (V N1 N2) (RC_1)	3.1%

CS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1)	38.9%
	(S) (V N1 N2 RC_1)	6.3%
N1) (N2 RC	(S V) (N1) (N2 RC_1)	4.2%
	(S V N1) (N2 RC_1)	3.2%
N1 N2) (RC	(S V) (N1 N2) (RC_1)	41.1%
	(S) (V N1 N2) (RC_1)	6.3%

High-attaching long RC

FS

Phrasing category	Phrasing pattern	
N1 N2 RC		0%
N1) (N2 RC		0%
N1 N2) (RC	(S V) (N1 N2) (RC_1 RC_2 RC_3)	94.7%
	(S) (V N1 N2) (RC_1 RC_2 RC_3)	3.2%
	(S V) (N1 N2) (RC_1) (RC_2 RC_3)	2.1%

CS

Phrasing category	Phrasing pattern	
N1 N2 RC		0%
N1) (N2 RC		0%
N1 N2) (RC	(S V) (N1 N2) (RC_1 RC_2 RC_3)	91.7%
	(S) (V N1 N2) (RC_1 RC_2 RC_3)	6.3%
	(S V N1 N2) (RC_1 RC_2 RC_3)	1%
	(S V) (N1 N2) (RC_1) (RC_2 RC_3)	1%

Low-attaching long RC

FS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1 RC_2 RC_3)	1.0%
N1) (N2 RC	(S V N1) (N2 RC_1 RC_2 RC_3)	4.2%
	(S V) (N1) (N2 RC_1 RC_2 RC_3)	2.1 %
	(S) (V N1) (N2 RC_1 RC_2 RC_3)	2.1%
	(S V N1) (N2) (RC_1 RC_2 RC_3)	2.1%
	(S V) (N1) (N2) (RC_1 RC_2 RC_3)	2.1%
	(S V N1) (N2 RC_1) (RC_2 RC_3)	1.0%
	(S) (V N1) (N2 RC_1) (RC_2 RC_3)	1.0%
	(S V) (N1) (N2 RC_1) (RC_2 RC_3)	1.0%
N1 N2) (RC	(S V) (N1 N2) (RC_1 RC_2 RC_3)	78.1%
	(S) (V N1 N2) (RC_1 RC_2 RC_3)	4.2%
	(S V N1 N2) (RC_1 RC_2 RC_3)	1.0%

CS

Phrasing category	Phrasing pattern	
N1 N2 RC	(S V) (N1 N2 RC_1 RC_2 RC_3)	1%
	(S V) (N1 N2 RC_1) (RC_2 RC_3)	2.1%
N1) (N2 RC	(S V) (N1) (N2 RC_1) (RC_2 RC_3)	1%
N1 N2) (RC	(S V) (N1 N2) (RC_1 RC_2 RC_3)	86.5%
	(S) (V N1 N2) (RC_1 RC_2 RC_3)	9.4%

Appendix C to Chapter 5: Perceived breaks over complete utterances.

All breaks perceived throughout the whole sentence are presented separately for FS and CS, each in the 4 experimental conditions created by the RC length (short/long RC) and attachment height (high/low) design. The following potential break points were considered: breaks after the subject, the matrix verb, after N1, after N2, and within the RC (after the subordinate verb). Breaks after the accusative marker *'et*, the preposition *šel* and after the second prosodic word within the RC were rare and therefore were set aside. The figures below report the amount of breaks perceived at any of the considered points. Figure 1 reports perceived breaks in the FS conditions; Figure 2 presents the perceived breaks in the CS conditions.

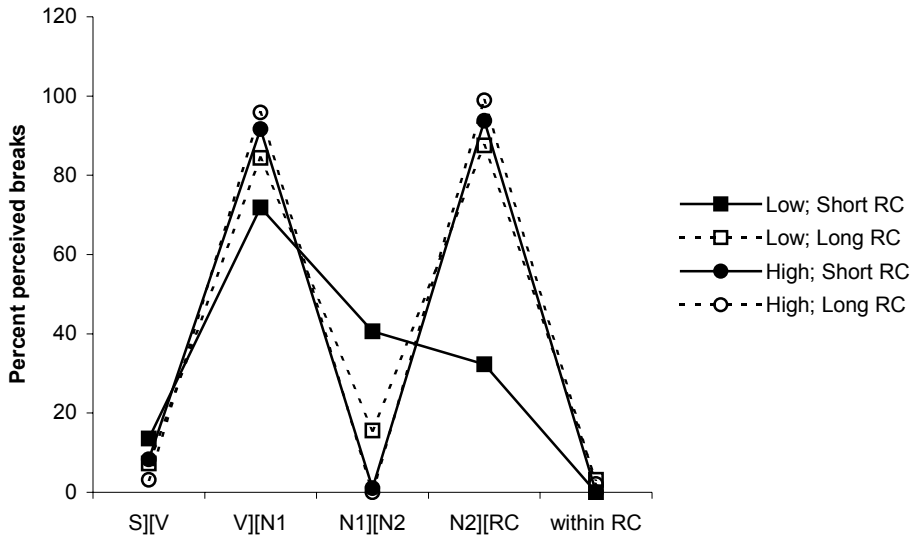


Figure 1: Perceived breaks in FS conditions, as a function of attachment height and RC length.

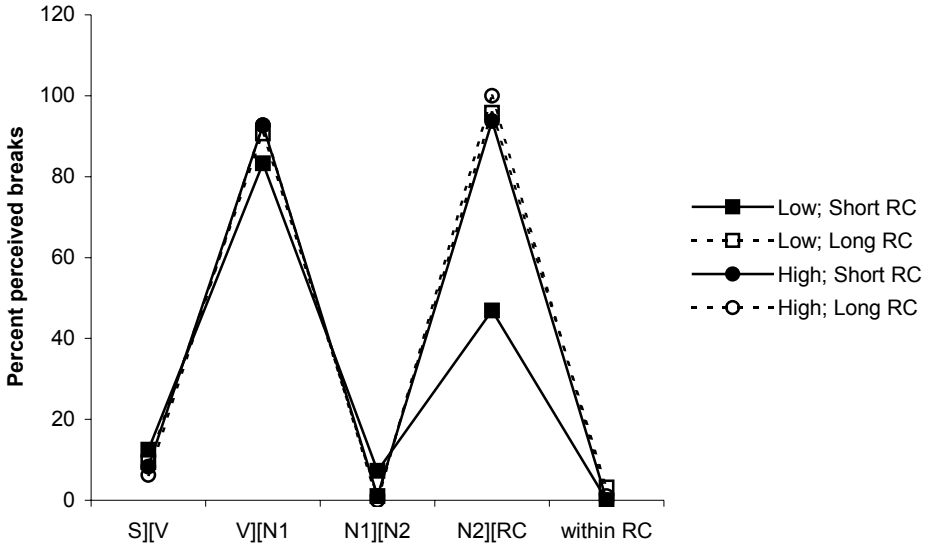


Figure 2: Perceived breaks in CS conditions, as a function of attachment height and RC length.

Note that there are extremely few breaks between the subject and the verb, and within the RC, in any of the conditions for either construction. In both the FS and CS utterances, a general pattern can be observed in three of the four conditions tested, namely the high attachment conditions and the low attachment long RC conditions. This pattern exhibits a break after the verb and an additional break following N2, with very few perceived breaks anywhere else in the sentence. The *N1 N2)(RC* pattern predominated over all alternative patterns in these three conditions.

The condition that varies greatly in terms of the location of breaks perceived over the whole sentence is the short RC attaching low. However, this condition varies in different ways for FS and for CS. In the FS, the low-attaching short RC shows fewer breaks after the verb, more breaks after N1, and fewer breaks after N2. This corresponds to the perceived phrasing analysis, which found a balanced choice between the phrasing

patterns *N1)(N2 RC* and *N1 N2)(RC*, and *N1 N2 RC*. CS utterances of a low-attaching short RC also showed variation from the predominating pattern, though only with respect to the position between N2 and the RC, where fewer breaks were exhibited compared with the three other conditions. This corresponds to the finding in the perceived phrasing analysis that there was a balanced choice between the *N1 N2)(RC* pattern and a pattern with no breaks within the N1 N2 RC sequence. Note that unlike FS, the CS data show no increase in breaks between N1 and N2 in the low attaching short RC condition, in accord with its assumed phonological status as a single phonological word.

CHAPTER 6

PERCEPTION OF SPOKEN SENTENCES [119]

6.1 Goals

The perception experiment reported in this chapter examined Hebrew native speakers' sensitivity to differences in the overt prosody of spoken sentences containing the ambiguous RC construction. There has been an increasing number of psycholinguistic studies over the last several years which focus on the interpretation of ambiguous sentences presented auditorily in perception tasks, in order to assess the contribution of (overt) prosody to the resolution of ambiguity for various constructions (e.g., Baltazani, 2002; Clifton et al., 2002, 2006, Hirschberg & Avesani, 1996; Kraljic & Brennan, 2005; Schafer, Speer & Warren, 2003; Snedeker & Trueswell, 2003, among others; see also seminal earlier studies by Lehiste, 1973 and Price et al., 1991). The listening experiment described in this chapter evaluates the contribution of overt prosody to the resolution of the RC attachment ambiguity in Hebrew.

The production study reported in Chapter 5 established an array of prosodic phrasings used in Hebrew. Three distinct phrasing patterns were found to be associated with the RC construction: one category of phrasing patterns had a break after N1; another category had a break after N2; the last phrasing category had no break in either position (neither after N1 nor after N2). The production study revealed the phrasing patterns used for the RC construction that are associated with each condition of the construction type x

[119] The research reported in this chapter was supported in part by a Graduate Research Grant from the Office of Research and Sponsored Programs of the CUNY Graduate Center.

RC length general design. Furthermore, since it presented sentences with an unambiguous intended meaning, it yielded data on the prosodic phrasing(s) associated with each attachment interpretation.

The first goal of the listening experiment was to explore whether the three distinct phrasing categories which were found in the production study result in an auditory perception task in a similar attachment interpretation to that associated with the corresponding meaning in the production study. Thus the listening study examines the interpretation of placement of prosodic breaks in different positions, under the assumption that the prosodic structure suggests a corresponding syntactic representation and is shaped by that structure, and thus a prosodic break is placed in a certain position in order to obey the ranking of the prosodic constraints of the language. Selkirk's (2000) formulation of the phonological constraints affecting prosodic phrasing (e.g., $Align_R$ *XP, Bin(MaP)*) suggests guidance to speakers as to placement of prosodic breaks in a sentence, but of course the constraints can be interpreted in the other direction as guidance to listeners on the syntactic interpretation of prosodic breaks they hear.

Thus, the corresponding working assumption is that a discontinuity in the prosody will cause perceivers to project a corresponding discontinuity in the syntax. Specifically, a prosodic break is associated with the end of a syntactic phrase. Three phrasing patterns observed in the production study were compared in the listening experiment. In Phrasing 1, a prosodic break occurs after N1 only; Phrasing 2 has a prosodic break after N2 with an earlier break after the verb; Phrasing 0 has no break either after N1 or after N2 but does have a break after the verb. These three phrasings are hypothesized to result in different attachment preferences for otherwise identical sentences. The three phrasing

patterns are outlined in (1)[120]. See Appendix A to this chapter for motivation of the placement of prosodic breaks earlier in the sentence.

(1) a. Phrasing 0: (S V) (N1 N2 RC)

b. Phrasing 1: (S V N1) (N2 RC)

c. Phrasing 2: (S V) (N1 N2) (RC)

Phrasing 0 and Phrasing 2 were used with sentences containing both an FS and a CS. Phrasing 1, on the other hand, was used with the FS only. It was decided that the two design factors (prosodic phrasing and construction type) would not be fully crossed because the prosodic pattern associated with Phrasing 1 is an intuitively unnatural pattern for sentences containing the CS. This is supported by the discussion on the status of the CS as a single prosodic unit (Chapter 2), and also by the findings from the production study that such a phrasing pattern is extremely rare for the CS. It was felt that presenting anomalous combinations of prosody and syntax might dull participants' sensitivity to prosody throughout the experiment. Further details on this aspect of the design of the listening study are included in §6.2.

The data from the production study allow us to make the following predictions regarding the interpretation of these three phrasing patterns. Since Phrasing 1 was only found in the low attachment productions, it is predicted to result in a low attachment interpretation of the RC: a break after N1 groups N2 and the RC into a single prosodic unit, with no prosodic break intervening between them, thus giving rise to a low

[120] The representation of the different prosodies does not include the accusative Case marker *'et* and the linking element *šel* (of) between N1 and N2 in the FS nominal, since they do not carry any prosodic weight by themselves and therefore are not relevant for a discussion of prosodic structure. It is assumed that both *'et* and *šel* cliticize onto the word that follows them (N1 and N2 respectively).

attachment interpretation of the RC.[121] Note that the N1)(N2 break is not dictated by the Selkirkian alignment constraint, and although it could be explained as a length-induced break, such a break would more likely be placed at the left edge of the RC. The N1)(N2 break is an interesting phrasing in light of the Rational Speaker Hypothesis (Clifton, Carlson & Frazier, 2006), which might be taken to predict that such non-Selkirkian breaks would not be interpreted by the parser as indicating a phrasal break and thus should not affect interpretation. It is of interest, therefore, to see the effect of Phrasing 1 on the interpretation of ambiguous sentences.

Phrasing 2, on the other hand, has been found to be the predominant pattern in the high attachment conditions, and was found to be only one of several phrasing patterns in low attachment of short RCs in the production study. Phrasing 2 was also found in the low-attaching long RC conditions of the production study; see discussion below. 60% of the occurrences of this phrasing pattern in the production study were associated with a high attachment interpretation; therefore it may be anticipated that there will be a tendency for sentences using this phrasing pattern to be perceived with high attachment interpretation rather than with low attachment interpretation in the listening study.

The interpretation of sentences using Phrasing 0, in which the whole complex nominal, including the RC, is parsed as a single prosodic unit, may perhaps be unaffected by prosodic structure, since there are no prosodic breaks in the crucial region of the sentence. Several possibilities can play a role in attachment decisions in a 'no break' phrasing: It is possible that the *absence* of a prosodic break in the critical region suggests a low attachment interpretation of the RC. Alternatively, prosody can be seen as neutral,

[121] Schafer (1997) found in her study of PP attachment ambiguity that attachment within a prosodic phrase is favored over attachment across prosodic boundaries.

with attachment preference affected by non-prosodic factors. For example, grouping of the RC with the complex nominal may allow for the locality of N2 to give preference to a low-attaching interpretation, as proposed by the principle of Late Closure: "if grammatically permissible, attach new items into the clause or phrase currently being processed; i.e., the phrase or clause postulated most recently" (Frazier, 1987:562) and Recency: "preferentially attach structures for incoming lexical items to structures built more recently" (Gibson et al., 1996:26). Finally, it is possible that the prosody-neutral Phrasing 0 results in no attachment preference. This phrasing pattern was found mostly in the low attachment conditions of the production study, suggesting that it is indeed associated with a low attachment interpretation of these sentences. However, it is predicted that the attachment preference will not be as low as that of Phrasing 1, since the presence of a prosodic cue (as in Phrasing 1) is predicted to have a stronger effect than the absence of prosodic cues (as in Phrasing 0), which is predicted to have a more subtle effect on attachment choice.

The listening study thus examines whether variation in the prosodic phrasing of a spoken target sentence, using the three phrasing patterns presented above, affects its interpretation in terms of attachment choice. A second goal is to examine whether prosodic phrasing affects interpretation of ambiguous productions of the RC construction in the same way regardless of other factors such as construction type and RC length. The experiment tests the extent to which prosody contributes to attachment interpretation by contrasting FS and CS: it examines whether sentences which are minimally different, i.e. one containing an FS and the other containing a CS, but which have the same prosodic patterns, are interpreted similarly in terms of RC attachment (e.g., whether sentences with

a prosodic break following N2 will be interpreted similarly, regardless of whether they contain an FS or a CS). To address this question, two phrasing patterns, Phrasing 0 and Phrasing 2, were each compared in sentences with FS and CS. Phrasing 1 was not compared between FS and CS, since the production study had found it to be an unnatural phrasing pattern for CS sentences (see above). The extent to which the contrast between the FS and CS plays a role when prosody is kept constant may help to evaluate whether the source of the difference in attachment preferences for these two complex nominals is prosody alone, or other, non-prosodic factors.

It was hoped that the experiment could contribute to evaluating the effect of prosody on attachment choices by examining whether sentences that contrast in RC length but are otherwise identical in terms of their prosodic structure result in similar attachment choices. It is not well established what length effects there are in how listeners respond to prosodic breaks; the only existing data for RC attachment in listening are in Fernández (2007), which found no difference in attachment due to RC length. An interesting proposal in the literature is the Rational Speaker Hypothesis (RSH) of Clifton et al. (2006), which predicts that a break before a long constituent (such as, in our case, a long RC) is likely to be interpreted by listeners as induced by length requirements and hence is less likely to be treated as a signal of syntactic structure, compared with a prosodic break before a short constituent. Thus in the case of a pre-RC break, the RSH would predict that a long RC will tend to evoke fewer high attachments than a short RC. It would have been interesting if the Hebrew data presented here could bear on this issue, but due to other design considerations the difference in RC length was tested only in the CS, which turns out to be unusually insensitive to length in general; see discussion below.

The motivation for the choice of RC length contrast in the CS is presented in Appendix B to this chapter. Phrasing 2 in the long and short versions is illustrated in (2a,b). In the notation used throughout this chapter, RC_1 RC_2 RC_3 represents the three prosodic words within the long RC: RC_1 represents the first prosodic word. RC_2 the second, and RC_3 the third, within the RC. RC_1 is also the whole RC, which is a single prosodic word, in the case of a short RC).

(2) a. CS: Phrasing 2-Short: (S V) (N1 N2) (RC_1)

b. CS: Phrasing 2-Long: (S V) (N1 N2) (RC_1 RC_2 RC_3)

As noted, the RC length contrast was examined with respect to CS sentences only. This was because there was no reason to assume that FS and CS will behave differently with respect to the RC length manipulation with the same prosody: the production data show that the difference between a short and long RC was almost identical in the CS (48.5% in low attachment, 6.3% in high attachment) and in the FS (51% in low attachment; 7.2% in high attachment), suggesting that if a contrast is found in the interpretation of sentences uttered with Phrasing 2 for long RCs compared with short RCs, it should be equally so for FS and CS. Excluding a length comparison in the FS items made it possible to limit the experiment to a relatively manageable size. Furthermore, since the CS was not used with Phrasing 1, as explained above, using the CS with Phrasing 2-Long had the advantage of creating a balanced number of occurrences for FS and CS throughout the experimental items; see further discussion in §6.2.

In sum, the listening experiment was designed to assess (a) whether varying the prosodic structure of a given sentence affects attachment choice; (b) whether keeping the prosodic structure constant results in a similar interpretation for the two different nominal constructions (CS and FS); and (c) whether the length (and consequently prosodic weight) of an RC affects the interpretation when both the prosodic structure and the construction type remain constant.

In preparing the stimuli for this study, an effort was made to ensure that the spoken sentences, though clearly intelligible, do not have exaggerated prosody but fall within the range of natural productions, thus making it possible to generalize the results of this experiment to spoken language spontaneously produced. (See Appendix C to Chapter 6 for a comparison of natural productions and semi-natural spoken sentences used in this experiment. Natural productions were elicited from a speaker using the Bradley et al.'s (2003) Post-to-Times technique (described in Chapter 5, §5.2.3), and served as a baseline for comparison with the semi-natural utterances produced by the same speaker as stimuli for the listening experiment.)

6.2 Method

6.2.1 Materials and design

Twenty-four experimental items were used in the listening study. These were the 24 items constructed for the set of experiments reported in this dissertation, as described in Chapter 4, and included the 8 items used in the production study (Chapter 5). See Appendix A-3 for a list of experimental items. Three design factors are manipulated in

this experiment: the first factor is prosodic phrasing pattern (Phrasing 0, Phrasing 1 and Phrasing 2), as specified in (1) above; the second factor is construction type (FS or CS); the third factor is RC length (short RC or long RC), as shown in (2).

The three factors are not fully crossed. As mentioned above, Phrasing 1 is common only in FS; it is highly uncommon in CS, therefore it was decided not to use it with the CS. In addition, as argued for above, no difference between FS and CS was expected to emerge from a comparison of Phrasing 2-Short and Phrasing 2-Long; therefore, it was decided to simplify the design by comparing Phrasing 2-Short and Phrasing 2-Long for CS only. Apart from this RC length comparison of Phrasing 2 for the CS, all remaining experimental items had an RC consisting of one prosodic word (a short RC). Table 6-1 presents a summary of the conditions tested in the listening experiment (empty cells represent conditions which were not tested). A description of three combined designs is presented towards the end of this section.

Table 6-1: Experimental conditions tested in the listening experiment, as a function of prosodic phrasing (Phrasing 0/ Phrasing 1/ Phrasing 2), construction type (CS/FS), and RC length (short/long RC). A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3).

		Phrasing 0	Phrasing 1	Phrasing 2
FS	Short RC	(S V) (N1 N2 RC)	(S V N1) (N2 RC)	(S V) (N1 N2) (RC)
CS	Short RC	(S V) (N1 N2 RC)		(S V) (N1 N2) (RC)
	Long RC			(S V) (N1 N2) (RC_1 RC_2 RC_3)

The choice of RC length for the target sentences was based on the naturalness of the resulting prosodic structure. If the length of the RC remains fixed at 1 prosodic word,

then all patterns of phrasing should be felicitous. From the finding that in Hebrew, the optimal length for a prosodic phrase is 2 prosodic words (Shaked, 2007; see also Chapter 5), it follows that an RC longer than 1 prosodic word, if used with all three alternative phrasing, would create for some phrasing patterns very long phrases, and thus would be more natural as a prosodic unit on its own. A short RC, on the other hand, allows more flexibility: it can form a unit on its own, be part of a 2-word phrase with N2, or form a three word phrase with both N1 and N2, all of which where evidences in the production study.[122] For these reasons, it was decided to use an RC of 1 prosodic word in the examination of the effect of overt prosody on interpretation of the target construction in the listening study; the one exception was the condition which specifically had a long RC of 3 prosodic words. Arguments for the choice of RC length for each of the three prosodies discussed above can be found in Appendix B to this chapter.

Three combined designs are created by the three factors of the experiment: Phrasing pattern (Phrasing 0, Phrasing 1 and Phrasing 2); construction type (CS or FS); and RC length (short or long). The first design compares three prosodic phrasings for the FS; the second design is a fully crossed 2x2 design of construction type and prosodic phrasing; and the third design contrasts RC length for the CS. Each of these three designs is described separately below.

The first design compared prosodic phrasing for FS sentences with RC length of 1 prosodic word, creating a single factor design with three levels. This made it possible to investigate the major phrasing patterns observed in the production study. Different predictions are made for each of these phrasing patterns, based on the production data (as

[122] For similar reasons but for a language with different prosodic properties, Lovrić (2003) used in his Croatian study a medium size RC length of 2 prosodic words.

described above): Phrasing 1 is predicted to result in low attachment, since it was only associated with low-attachment productions of this construction; Phrasing 2 is predicted to be associated with high attachment, due to its overwhelming use in the high attachment productions; in Phrasing 0, the absence of prosodic structure may give an opportunity for a syntactic recency effect to emerge, thus potentially driving attachment low. The three conditions contrasted in the first design are presented in Table 6-2.

Table 6-2: Comparison of FS utterances produced with three phrasing patterns: Phrasing 0, Phrasing 1 and Phrasing 2. RC contains 1 prosodic word uniformly.

	FS short
Phrasing 0	(S V) (N1 N2 RC)
Phrasing 1	(S V N1) (N2 RC)
Phrasing 2	(S V) (N1 N2) (RC)

Design 2 compared different prosodic phrasings for the two constructions studied in this dissertation. Each of the two constructions (CS, FS) was uttered with the two phrasing patterns that are grammatical for both. The two phrasings, Phrasing 2 and Phrasing 0, both have a break after the verb; they differ by an additional prosodic boundary after N2 in Phrasing 2 and its absence in Phrasing 0. (See Appendix A to chapter 6 for motivation of a break after the verb in Phrasing 2.) RC length is set to 1 prosodic word, uniformly. (See arguments for choice of RC length above.) Thus the second design uses two of the three cells that were already part of Design 1 (the FS conditions), and adds two further cells for the CS conditions. This design allows us to look at the interpretive consequences of different prosodic phrasings, and to assess

whether those interpretive consequences are independent of construction-type differences. A summary of the 4 conditions created by the second design is presented in Table 6-3 below.

Table 6-3: A fully crossed comparison of construction type (FS or CS) and phrasing patterns (Phrasing 0, Phrasing 1 and Phrasing 2), with RC uniformly 1 prosodic word.

	CS short	FS Short
Phrasing 0	(S V) (N1 N2 RC)	(S V) (N1 N2 RC)
Phrasing 2	(S V) (N1 N2) (RC)	(S V) (N1 N2) (RC)

The third design compared two lengths of RC in CS sentences. In Designs 1 and 2, RC length has been set at 1 prosodic word. To find out whether the interpretive consequences of a prosodic phrasing are constant when constituent weight changes, the design adds one cell (not yet considered in the previous two designs) in which RC length is set to 3 prosodic words. This single factor design has two levels, as shown in Table 6-4.

Table 6-4: A comparison of RC length (short and long) for CS utterances produced with Phrasing 2. A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3).

	CS
Phrasing 2, Short	(S V N1) (N2 RC)
Phrasing 2, Long	(S V) (N1 N2) (RC_1 RC_2 RC_3)

Note that it is inescapable, given these various design decisions, that Short-RC presentations stand in 5:1 ratio with Long-RC presentations, within the set of target constructions. This bias is counterbalanced by filler materials, as explained below.

The experimental design is thus a [(2 x 3 x 2) – 6)] design. The underlying design varies phrasing pattern (Phrasing 0, Phrasing 1, Phrasing 2), construction type (FS, CS), and RC length (short, long), as table 6-5 shows. Six cells in Table 4 are omitted on grounds of grammaticality and/or feasibility (cells shaded dark gray in table below) and economy (cell shaded light gray).

Table 6-5: Factorial design of the listening experiment varying phrasing pattern, construction type and RC length, with conditions omitted on feasibility/grammaticality grounds shaded in dark gray, and a condition omitted on economy grounds shaded in light gray. A short RC is represented in the notation as (RC) while a long RC is represented as (RC_1 RC_2 RC_3).

	Free State		Construct State	
	Short	Long	Short	Long
Phrasing 0	(S V) (N1 N2 RC) Grammatical and 2-3 is feasible	(S V) (N1 N2 RC_1 RC_2 RC_3) Grammatical but 2-5 not feasible	(S V) (N1 N2 RC) Grammatical and 2-3 is feasible	(S V) (N1 N2 RC_1 RC_2 RC_3) Grammatical but 2-5 not feasible
Phrasing 1	(S V N1) (N2 RC) Grammatical and 3-2 is feasible	(S V N1) (N2 RC_1 RC_2 RC_3) Grammatical but 3-4 is not feasible	(S V N1) (N2 RC) Ungrammatical	(S V N1) (N2 RC_1 RC_2 RC_3) Ungrammatical
Phrasing 2	(S V) (N1 N2) (RC) Grammatical 2-2-1 is feasible	(S V) (N1 N2) (RC_1 RC_2 RC_3) Grammatical 2-2-3 is feasible	(S V) (N1 N2) (RC) Grammatical and 2-2-1 is feasible	(S V) (N1 N2) (RC_1 RC_2 RC_3) Grammatical and 2-2-3 is feasible

Feasibility becomes an issue because the purposes of the experiment dictate the avoidance of any phrasing breaks within the target construction (i.e., within N1 N2 RC or

N1 N2 RC_1 RC_2 RC_3) other than those specified by the contrasts of interest. In particular, a break within a long RC makes the initial sub-part of the RC short enough to attach to previous material in the sentence, and thus might neutralize the short/long contrast; therefore, an RC-internal break is to be avoided. An RC internal break was indeed found to be rare in the production data (see Chapter 5). In order to avoid an RC-internal break, Phrasing 2 was chosen as the prosody with which RC length can be contrasted. Absent an extra phrasing break, Prosodies 0 and 1 are ruled out by feasibility.

The 6 different conditions created by the partial crossing of the three factors above are summarized in (3):

(3)

Condition 1:	FS short, Phrasing 0	(S V) (N1 N2 RC)
Condition 2:	FS short, Phrasing 1	(S V N1) (N2 RC)
Condition 3:	FS short, Phrasing 2	(S V) (N1 N2) (RC)
Condition 4:	CS short, Phrasing 0	(S V) (N1 N2 RC)
Condition 5:	CS short, Phrasing 2	(S V) (N1 N2) (RC)
Condition 6:	CS long, Phrasing 2	(S V) (N1 N2) (RC_1 RC_2 RC_3)

Each of the 24 experimental items had 6 variants, corresponding to the 6 experimental conditions created by the factorial combination in (3) above. Each participant heard one variant of each item; thus the 6 variants of each item were compared across participants. Each of the 6 questionnaire lists had 4 items in each condition, for a total of 24 experimental items per list.

In addition to the 24 experimental items, the listening experiment included 72 filler sentences, giving a 1:3 ratio of targets to filler items. Filler items were unambiguous sentences, and were identical across versions of the questionnaire. They were used to disguise the phenomena under investigation and to minimize the chance of participants identifying the experimental items and/or the purpose of the experiment. The fillers were

also used to filter out participants who were not attentive to the task, as revealed by inaccuracy of their responses to these items. Participants who made more than 15% errors on unambiguous filler items were not included in the final analysis of the data.

Similarly to the experimental items, filler items included some short RCs and some long RCs, and some filler sentences contained FS nominals while other contained CS nominals. However, these design factors were not combined together in filler sentences: a sentence containing a complex nominal did not have an RC modifying one of the nouns within the complex nominal. In addition, some filler sentences were constructed as simple sentences, though they were long enough not to stand out as superficially different from the rest of the fillers or experimental items. For filler sentences that included complex nominals, the following comprehension question did not relate to the nouns in the complex nominal, in order to avoid biasing participants to focus attention on the complex nominal in expectation of a question about it. For the same reason, for filler sentences that included an RC, the related question referred to the content of the RC in only half the sentences. The complete list of filler items, including a description of the different types of filler items and their distribution, is available in Appendix A-5.

As explained above, the design of the experiment yields an uneven number of short and long RC sentences in the experimental items, with a 5:1 ratio of short to long RC. This bias was matched by a 1:5 ratio of short to long RCs in the 24 filler items that had an unambiguously attached RC (24 of the 76 filler items). This created overall a balanced number of short and long RCs in the filler and experimental items combined.

Filler items included variations of the prosodic patterns used in the experiment as a whole. Some filler items were kept short, so that the experimental items that are short would not stand out as the shortest items in the experimental list. It was also ensured that the set of filler items had a proportion of breaks after N1 (similar to Phrasing 1) that matched the experimental items, such that the phrasing pattern corresponding to Phrasing 1 did not stand out amid the rest of the items in the questionnaire. This was achieved by the use of phrasal verbs in the filler items; as discussed in chapter 4, phrasal verbs were not used in experimental items.

The fillers were recorded in varied prosodic patterns, to match the different prosodies manipulated in the experimental items; however, in contrast to the experimental items, the prosodic patterns of the filler items do not affect the response to the binary choice, since they were all unambiguous. In most cases, each filler item was given its most natural prosody according to the experimenter's intuition. However, under the assumption that each of the 6 filler-types from which the filler items were constructed (see Appendix A-5) may have a default prosodic structure, it follows that there will be little variation in the prosodic patterns introduced by filler items. In order to include prosodic patterns similar to those of the experimental items, it was necessary to vary from the 'default' prosody for some filler items and use prosodies that are acceptable but less commonly used. As the production study showed, a given sentence may have several prosodic patterns, even when the same meaning is being conveyed; therefore, this decision did not jeopardize the naturalness of the utterances. It should be stressed that despite these manipulations, all filler items ended up having a natural prosody, as was judged by an expert listener (see below).

The experiment was run on a computer using DMDX display software. Each of the 6 lists of the experiment began with 3 overt practice items and 4 covert practice items (see Appendix A-5), which were uniform in all versions and appeared in the same order. Practice items were followed by 4 scrambled blocks. The items within each block appeared in individually-randomized order, and the order of the blocks themselves was random as well. Each block in itself was a mini-version of the experiment: it contained 1 item per condition (6 experimental items; 3 with animate complex nominals and 3 with inanimate complex nominals) and 16 filler items (varying in filler-type and prosodic phrasing), for a total of 22 items per block. The complete set of 4 scrambled blocks was then followed by 4 additional filler items, which ensured that the last item encountered by participants was not an experimental item. These were not part of the pseudo-randomization and appeared as the same last 4 items in all versions of the experiment. Thus participants heard 99 sentences during the experiment: 3 overt practice items, 24 experimental items (6 in each block) and 72 filler items (16 in each block, 4 as covert practice items and 4 as list-end protectors).

Since there were 4 scrambled blocks, with 1 experimental item per condition in each block, it follows that each of the 6 lists included 4 experimental items in each condition, for a total of 24 items. These quadruples were displayed with a different condition in each list, such that the total number of experimental sentences over all questionnaire versions was 144. The distribution of experimental items across conditions was such that each quadruple of experimental items contained 2 animate and 2 inanimate

items, and was presented in a different condition in each of the 6 lists of the experiment.[123]

In addition to experimental items, each of the 4 blocks included 16 filler items. Of the 16 fillers, there were between 2 and 3 of each filler-type. In each block, in half of the filler items answer (a) (which appeared on the right of the screen) was correct and in the other half answer (b) (which appeared on the left of the screen) was correct. Each block had the same type of prosodic phrasing distribution.

6.2.2 Recording and analysis of auditory stimuli

To prepare the auditory questionnaire, a trained Hebrew native female speaker recorded each item in the 6 conditions. Each sentence was presented separately to the speaker, printed in the middle of a page; the required positions of breaks were marked by large spaces between phrases. The example in (4) shows how each sentence was displayed to the trained speaker (but without the English gloss in (4)). Appendix A includes the prosodic patterns used for experimental (A-3) and for filler items (A-5).

(4)

ha-'ohadim he'ericu	'et ha-me'amen šel ha-mit'agref	še-paraš
the-fans admired	ACC the coach of the-wrestler	who-retired

האוהדים העריצו	את המאמן של המתאגרף	שפרש

The trained speaker read each sentence for recording one or more times, until both the author and the speaker were satisfied that the recording matched the intended prosody

[123] Quadruple A included items 1, 2, 13 and 14; B included items 3, 4, 15 and 16; c included items 5, 6, 17 and 18; D included items 7, 8, 19 and 20; E included items 9, 10, 21 and 22 and F included items 11, 12, 23 and 24.

and was natural. The speaker was instructed to pronounce each sentence as naturally as possible and not to exaggerate the pronunciation or the prosodic cues. The speaker was given examples of each of the prosodic patterns from the production study data and from a recording by the author, in order to establish what constituted a natural utterance of each prosodic pattern. In 6 recordings (out of a total of 144 experimental sentences), a falling tone was observed in the initial recordings, although a rising tone was expected. These 6 experimental sentences were re-recorded such that all phrasal breaks were uniformly marked by a rising tone. The newly recorded items were examined by the experimenter and were confirmed to exhibit a rising tone in the intended positions. They were then given to the trained judge who found them to be natural and clear, as were the rest of the original recorded materials.

Recording took place in a quiet laboratory at the CUNY Graduate Center. All sentences were recorded with the same microphone and volume level. Recording proceeded in 4 separate sessions. The 6 variants of each item were recorded in a single session, to allow for uniformity across conditions. Each session also included recording of filler items, for similar reasons. In each session, 6 experimental items (each in 6 variants), and 18 filler items were recorded, [124] for a total of 54 individual sentences per recording session. The recordings were pseudo-randomised such that each of the 4 design blocks in the questionnaire included a few items from each recording session, thus ensuring that the voice quality of the speaker and the recording quality were homogeneous throughout the experiment (see §6.2.1 for details on the design blocks).

[124] The 6 experimental items included 3 items with animate complex nominals and 3 items with inanimate complex nominals; the 18 filler items included 3 items from each filler type. See appendix A-5 for details on the various filler types used in this study.

The recorded stimuli were acoustically analyzed in the same fashion as were the recordings from the production study (Chapter 5), in order to establish that the sentences were uttered with the intended prosodies, with prosodic breaks located according to the design plan for each variant. If indeed the recordings follow the design plan, we expected to find in this analysis that for Phrasing 0 sentences, there was no lengthening of either N1 or N2, corresponding to no break after either nouns; for Phrasing 1, we expected N1 to be longer in duration compared to its duration measurements in the other conditions; and for Phrasing 2, we expected to find longer durations of N2, as compared to N2 duration in the other conditions, corresponding to a prosodic break after N2 in this condition.

Figure 6-1 below shows mean duration measurements for N1; Figure 6-2 shows the duration means for N2.

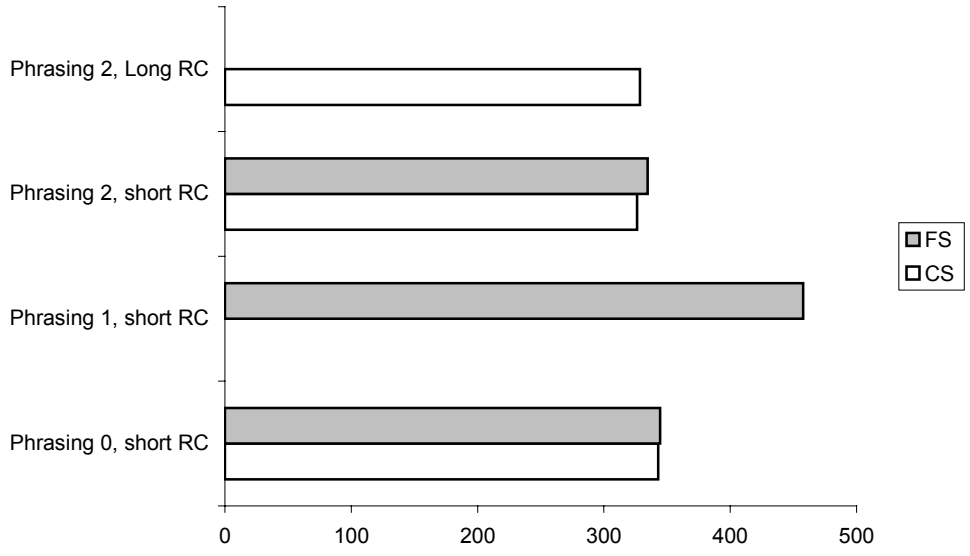


Figure 6-1: Mean duration for the N1 region in each of the 6 experimental conditions tested in the listening study.

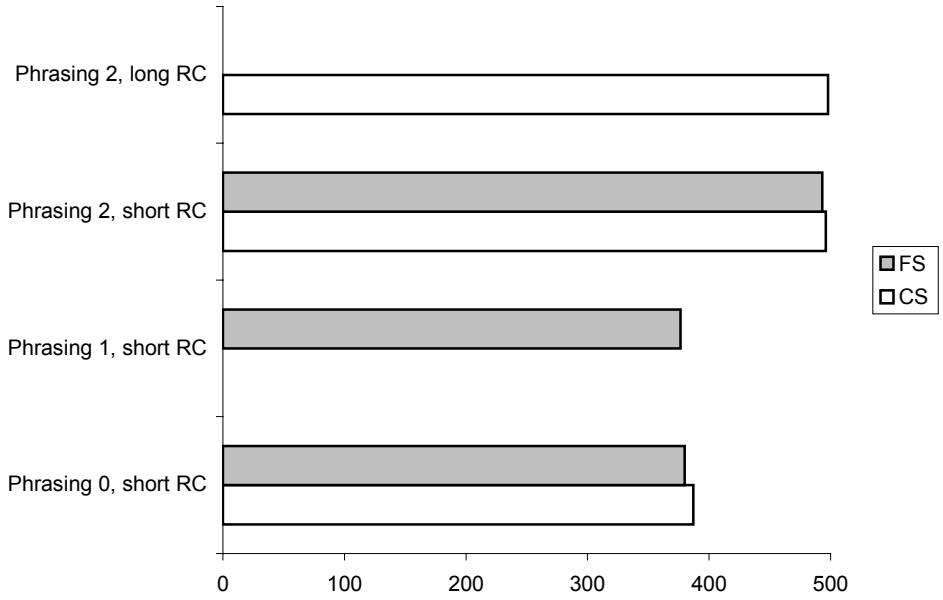


Figure 6-2: Mean duration for the N2 region in each of the 6 experimental conditions tested in the listening study.

This analysis shows that N1 duration was longer for Phrasing 1 (exhibited as condition 2 in (3) above) as compared to all other conditions combined, as expected: N1 was 122ms longer in condition 2, which exhibited Phrasing 1 (the N1 segment measured 458ms in condition 2, compared with 335ms on average in the remaining conditions). An omnibus ANOVA showed that at least one of the 6 conditions differed significantly from others ($F_2(5,115)=65.40$, $p<.001$), and a post-hoc planned comparison confirmed that condition 5 is significantly longer than the average of the 5 remaining conditions ($F_2(1,23)=199.24$, $p<.001$). Note than N1 duration did not vary greatly in the remaining conditions, whether it was part of an FS or a CS ($F_2(4,92)=2.06$, $p=.093$), suggesting that the stress reduction that the N1 of CS undergoes does not affect its length.[125] For N2 duration, an omnibus

[125] This is a welcome result, since it supports a straightforward comparison of the N1 segment between FS and CS in the production study (Chapter 5); it is not surprising, since the materials for the production study were specifically chosen such that the CS (bound) form does not undergo any morpho-phonological

analysis of variance confirmed that some of the conditions differed from each other in duration ($F_2(5,115)=121.52$, $p<.001$) and a post-hoc analysis showed that duration was longer in the conditions where a phrasal break following it was indicated (conditions 3, 5 and 6, where Phrasing 2 was instantiated, averaged 496ms), as opposed to its length in conditions where a break was not called for (N2 duration in conditions 1, 2 and 4, where Phrasing 0 and Phrasing 1 were instantiated, averaged 381ms) ($F_2(1,23)=396.07$, $p<.001$). The duration of N2 increased by an averge of 115 ms before a phrasal break. All ANOVAs for the listening stimuli are in Appendix D-2. With respect to durations, the recordings thus fully matched the conditions intended. Furthermore, the trained speaker produced increased durations of similar extent on N1 (122ms) and on N2 (115ms) in realizing a phrasal break. It was consequently decided that it was not necessary to perform an additional analysis of pitch accent, which would have been time consuming. Instead, the experimenter went over the recordings and determined by ear that each intended phrasal break was marked not only by lengthening but also by a rising tone.

In addition to establishing prosodic break positions, it was important to ensure that the recordings of experimental items were natural sounding in terms of their prosody. Establishing that the recorded materials are not exaggerated is necessary in order to claim that listeners are sensitive to prosodic structure. 'Natural prosody' refers to fluent speech that can be observed in colloquial, everyday communications of native Hebrew speakers (the register used in informal or unplanned speech such as a chat) and is not considered forced, exaggerated or carefully planned. Two native speakers (the experimenter and a trained judge) assessed the recordings and noted any unnatural or exaggerated utterances.

---

alternations as a result of stress reduction (see, e.g., Chapter 2, §2.3), which may have changed its overall length.

The trained judge was the same individual who supplied perceived phrasing judgments on the materials from the production experiment (Chapter 5), and therefore was thoroughly familiar with the materials. All recordings used in the listening study were judged to be clear, unexaggerated and natural to the native ear. They are posted on the author's website (https://wfs.gc.cuny.edu/ashaked/www/index.html). In addition, the recordings of the stimuli were compared with comparable sentences produced by the same speaker using the Post-to-Times protocol (similar to the one used in the production study reported in Chapter 5). An instrumental analysis of N1 and N2 duration of stimuli sentences where lengthening was expected found that all durations in the stimuli fall within the range of durations in Post-to-Times productions of comparable sentences by the same speaker (stimuli recording which had a break at a certain location were compared to those productions in which a break was perceived to occur at that location). See Appendix C to this chapter for details on this comparison, which contributes to an assessment of the naturalness of the stimuli for the listening study.

6.2.3 Procedure

Participants were tested individually, in a quiet room within the facilities of the Linguistics department at Tel Aviv University. The experiment was run by two paid experimenters, who were recruited and trained by the author of this dissertation in Israel during January 2006.

Participants were given a printed sheet of instructions, which prepared them for the task they were about to perform on the computer. They then listened through headphones to utterances played on a Dell laptop. Before the experiment began,

participants saw a short version of the instructions once again, this time on the computer screen. The printed instructions as well as the short version of these instructions are presented in Appendix B-2. Participants then had a short practice session, in which they had a chance to practice the procedure (see Appendix A-5 for practice items), after which they were able to ask any questions they might have concerning the procedure. It was made clear that once the experiment began, it should proceed without interruption. Although participants were encouraged to respond intuitively, the instructions did not specifically require speeded responses.

Participants listened to the spoken sentences, each followed by a comprehension question displayed on the screen. Their task was to choose between two possible answers to the comprehension question. For ambiguous items, the two answer choices corresponded to the two interpretations of the sentence. Each item was preceded by a visually-presented 'play' symbol "►" in the middle of the screen. To begin the task for an item, participants had to respond by pressing the space bar. Immediately after pressing the space bar, they heard a sentence. While the sentence was playing, the computer screen remained blank. Participants had a choice of pressing a key on the keyboard marked 'replay' in order to hear the same sentence again.[126] After hearing the sentence

[126] A second hearing gives subjects an opportunity to listen to sentences again in case they were distracted the first time, thus reducing the risk of their making a random choice between the two answers. Since the experiment did not examine reaction time, it was not crucial to analyze first listening only; furthermore, responses of participants who had a second listening were not expected to vary in any systematic way from responses of other participants who only had a first listening. The DMDX software kept track of the number of times each sentence was played: participants replayed sentences in 9% of experimental items and 8% of filler items; 6% of participants (7 participants) used the 'replay' option in more than ¼ of experimental items (similar in filler items), while 30% of participants did not use that option at all in experimental items (23% in filler items). An analysis of the data excluding the 7 participants who replayed sentences in more than ¼ of experimental items showed almost no variation from the means which include their data, suggesting that this aspect of the procedure did not interfere with the results of the experiment. An analysis of the filler error rate for these 7 participants showed that they averaged 4.2% errors on fillers; this low error rate, compared to the error rate cutoff of 15%, suggests that they paid careful attention and performed well on the task.

(once or twice), participants pressed the space bar and a question about the item appeared on the screen, with 2 answer choices, one on the right side of the screen and the other on the left side of the screen. For experimental sentences, the question queried the RC-attachment choice; the two choices offered were N1 and N2. For filler items, the question was about an action described in the filler sentence and the options were two nouns that appeared in the sentence (only one choice was correct). The order of presentation of the two answer choices matched their order of occurrence in the sentence. Participants pressed the right 'shift' button if the answer they chose was on the right of the screen, and the left 'shift' button if they chose the answer on the left of the screen. Once they chose an answer, a new screen appeared with a 'play' symbol, indicating the beginning of a new item. This procedure allowed participants to pace themselves, since they controlled when the sentence was played and when the question was presented (within reasonable time constraints; a timeout of 9 seconds for question presentation and for question answering was inserted into the program with feedback so that participants would stay alert).

The experiment took an average of 45 minutes to complete, including instructions, practice session and the list of items. A graduate student in linguistics who is a native speaker of Hebrew verified that the procedure was clear and easy to follow in all 6 lists, before collection of data began.

6.2.4 Participants

One hundred and twenty students at Tel Aviv University participated in the listening experiment successfully. They were all native speakers of Hebrew, as ensured

by a background questionnaire, similar to the one used in the production experiment (see Appendix C for the background questionnaire). Participants were naïve with respect to the purposes of the experiment. Each participant received 30 New Israeli Shekels (NIS), equivalent to approximately $6 at the time, for participation. This is the usual rate for 1 hour of participation in psychology experiments in Israel.

Twelve participants were excluded from the analysis of the data because they failed the error rate criterion: they answered incorrectly on 15% or more of the filler items, suggesting that they did not pay attention to the task. Also, two participants were excluded from the analysis of the experiment because they were not native speakers of Hebrew.

6.2.5 Data measurement and assessment

Responses to experimental items were coded in terms of attachment site (high attachment or low attachment), as implied by the choice between N1 (the right-side answer) and N2 (the left-side answer).[127] Summary values were cast in terms of %N1 indicating high attachment interpretation.

The analysis of results follows the three separate designs presented in §6.2.1 above. Design 1 compares, within the FS, the three different prosodies: Phrasing 0, Phrasing 1 and Phrasing 2. This comparison shows the effect of prosodic structure on the interpretation of a single sentence. Design 2 compares the two construction types (FS and CS) each in two different prosodic patterns: Phrasing 0 and Phrasing 2. This comparison examines whether identical prosodic phrasing patterns affect the two constructions in the

[127] The Hebrew orthography follows a right to left directionality and therefore N1 (the first noun in the sequence) appears to the right of N2 in the orthography.

same way. The comparison enables us to assess whether prosodic factors are the major factor affecting attachment choices for these comparable constructions, or whether there are other, non-prosodic factors interacting with prosodic phrasing. Finally, design 3 compares, within the CS, the effect of RC length on the interpretation of a sentence, when prosody remains constant. This analysis examines whether a break before a short RC, compared to a similar break before a long RC, results in a similar or different interpretation of the RC.

6.3 Analysis and results

Summary data for the three designs of the experiment are presented separately below. (All ANOVAs for the listening experiment are provided in Appendix D-2.) Figure 6-3 presents summary data for the listening of FS sentences in varying phrasing patterns (Phrasing 0, Phrasing 1 and Phrasing 2). RC is uniformly set to 1 prosodic word.[128]

[128] Percentages for all data analyses are calculated over the total of 120 responses per condition; timeout responses (missing data points) reflect 1.6% of total responses.

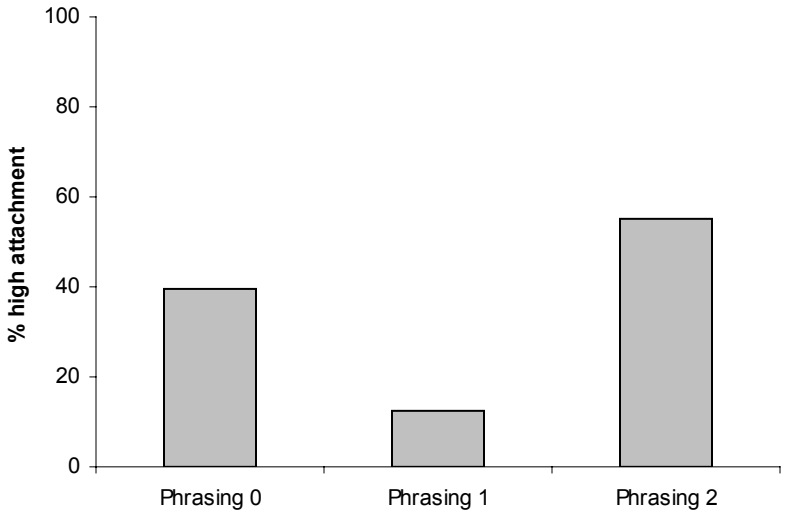


Figure 6-3: Design 1, FS Listening data in terms of % high attachment, for each of the 3 phrasing patterns: Phrasing 0, Phrasing 1 and Phrasing 2. RC contains 1 prosodic word, uniformly.

A comparison of the three phrasing patterns within the FS with short RC (design 1) reveals a striking difference in the interpretation of the three prosodic patterns, with a main effect observed for prosody (F_1 (2,228)=131.37, p<.001; F_2 (2,36)=68.45, p<.001). The likelihood of high attachment is highest with Phrasing 2, less likely with Phrasing 0 and least likely with Phrasing 1, in line with predictions outlined above. Phrasing 2 yielded significantly higher attachment than Phrasing 0 (F_1 (1,114)=30.37, p<.001; F_2 (1,18)=23.85, p<.001); Phrasing 0 yielded significantly higher attachment than Phrasing 1 (F_1 (1,114)=124.22, p<.001; F_2 (1,18)=53.74, p<.001); and therefore, Phrasing 2 yielded significantly higher attachment than Phrasing 1. This confirms our predictions that hearers preferred to attach the RC low when the sentence had a break after N1 (12.5%), and were more likely to interpret the RC as attaching higher when the break occurred after N2 (55.1%). With respect to Phrasing 0, the prosodic structure does not

provide any guidance with respect to the RC interpretation. The data show that attachment interpretation is likely to be low (39.4% high attachment), though significantly higher than that of Phrasing 1. This suggests that while other factors (e.g., Late Closure/Recency or the absence of a prosodic break guiding attachment interpretation) may encourage the low interpretation, their effects are not as categorical as a break after N1 (Phrasing 1), which clearly results in a low attachment interpretation, leaving little chance for the alternative attachment choice.

Listening data for the analysis of construction type (FS/CS) and phrasing pattern (Phrasing 0/ Phrasing 2) are presented in Figure 6-4. RC length is fixed at 1 prosodic word.

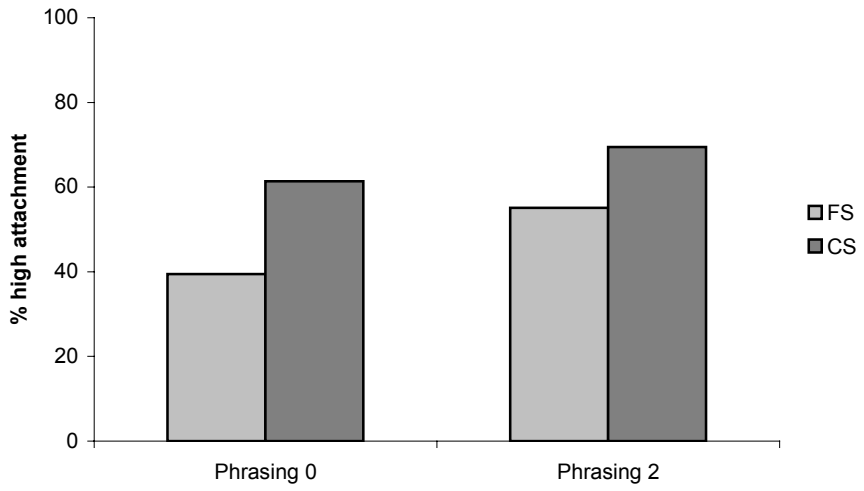


Figure 6-4: Design 2, Listening data in terms of % high attachment, for CS and FS in two phrasing patterns: Prosody 0 and Prosody 2. RC length is fixed at 1 prosodic word.

A comparison of construction type (FS/CS) and phrasing pattern (Phrasing 0/ Phrasing 2), with a short RC (design 2), reveals an effect for construction type (F_1 (1,114)=67.44, p<.001;F_2 (1,18)=45.68, p<.001), an effect for prosody (F_1 (1,114)=38.90, p<.001; F_2

(1,18)=37.42, $p<.001$) and a marginal interaction (F_1 (1,114)=3.95, $p=.05$; F_2 (1,18)=3.41, p= .0813, $.05<p<.1$). There is a striking difference between the FS and CS in that the CS is consistently interpreted higher than the FS (overall mean for FS and CS is 47.3% and 65.4%, respectively), for which some explanation may be sought. There is also a hint that the CS is less sensitive to the prosodic phrasing differnce than the FS is. To explore this further, a sub-analysis looked at the difference between Phrasing 0 and Phrasing 2 for each of FS and CS separately.

The analysis shows a greater likelihood, both in FS and CS sentences, for Phrasing 2 to be perceived as a signal of high attachment than Phrasing 0. For the FS, attachment was significantly higher for Phrasing 2 than Phrasing 0 (F_1 (1,114)=30.37, $p<.001$; F_2 (1,18)=23.85, $p<.001$). Similarly, attachment was significantly higher for Phrasing 2 than Phrasing 0 for sentences with CS nominals (F_1 (1,114)=10.10, $p<.005$; F_2 (1,18)=8.53, $p<.01$). The data thus confirm the prediction that Phrasing 2 sentences are interpreted as attaching higher than Phrasing 0 sentences, not only for FS sentences, as concluded above (in design 1), but also for CS sentences. The generality of the effect of prosodic phrasing on attachment preference observed for the FS above is supported by the finding of a similar prosodic effect observed in the CS. It should be noted, though, that while the difference in interpretation of CS with Phrasing 0 and Phrasing 2 (61.4% and 69.4%, respectively) was significant, the numerical difference between the two conditions was smaller than in the FS (39.4% and 55.1%, respectively), and the effect of prosody, though present, was reduced in magnitude compared to the interpretation of these two comparable conditions in the FS. Furthermore, the prediction of a low attachment interpretation for Phrasing 0 was not borne out in the CS data. Attachment

interpretation was high (61.4%), contrary to expectations. It appears that the reduced effect of prosodic phrasing in the CS is due to this data point. The finding that attachment varies between the FS and CS when prosody remains constant suggests the existence of other, non-prosodic factors interacting with prosodic phrasing in the data reported here. In particular, the effect of prosody is reduced in the CS cases, compared to the effect of prosody in the FS cases, suggesting that factors internal to the CS increase the likelihood of high attachment in the CS, and override the effect of prosody.

Finally, we turn to the analysis of the third design. Data on the comparison of the effect of RC length on the perception of prosodic phrasing (design 3) is presented in Figure 6-5. The data compared the CS with Phrasing 2 in two RC lengths: short RC (1 prosodic word) and long RC (3 prosodic words) and found that a short RC was interpreted as 69.4% high attachment and a long RC was interpreted as 69.2% high attachment.

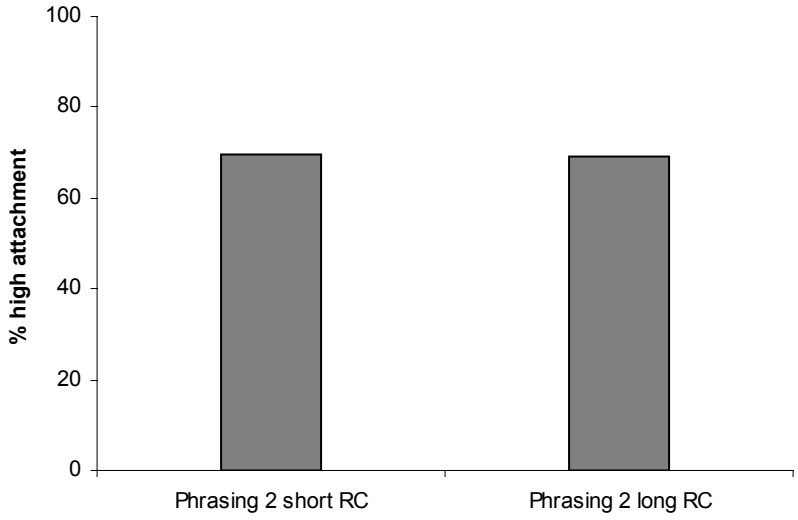


Figure 6-5: Listening data in terms of % high attachment, for CS sentences with Phrasing 2 in two RC lengths: Short RC (1 prosodic word) and long RC (3 prosodic words).

When comparing short and long RCs using a single prosody (Phrasing 2) for the CS (design 3), no difference was found (F_1 <1; F_2 <1), suggesting that perceivers interpret a pre-RC break in the same way, as signalling high attachment, regardless of the length of the RC. The lack of difference in interpretation preference for short and long RC is interesting, since it suggests that the added informativeness of the RC does not affect attachment interpretation, thus contradicting the hypothesis on 'information load' proposed by Hemforth et al. (2000). This hypothesis offers an alternative to the prosodic approach, e.g., it offers a pragmatic account of the modifier attachee length effect (in our discussion, a long RC): A long RC contains more information, which means more processing effort. The enhanced processing effort of a longer RC must be justified, and is justified if the long RC modifies a more central element in the proposition (an argument of the verb). According to this hypothesis, a more informative constituent would have a higher tendency to attach to a more prominent element (N1, an argument of the main verb, as opposed to N2, a modifier of N1). The informational account of modifier attachee length attributes high attachment to the richness of the attachee rather than to prosodic weight. This hypothesis is supported by data: when lengthening the modifier (a PP, in Hemforth et al's (2006) study of German and French) without adding any semantic information (e.g., *Au* vs. *Au-Winterach*), there was no effect of length. When adding an adjective (e.g., *Au* vs. *dem schoenen Au* 'beautiful Au'), there was higher tendency for high attachment. Note that the long RC sentences of this experiment were carefully designed such that little additional content was added to the short RC version, in order to hold informativeness as constant as possible (see Chapter 4 for more details); therefore it is unlikely that this explanation holds for the data presented here. The results for the CS

show that indeed the added content of the long RC condition does not affect the preferred attachment of the RC, compared with the short version of the RC. While it was found that prosodic phrasing manipulations affect attachment interpretation to some extent (though the effect was less pronounced than that observed in the FS), as discussed above, length manipulations did not result in any variation in attachment interpretation in the CS. Nonetheless, since length variation was unfortunately not tested in the FS in the listening experiment, and since the CS seems to exhibit inherent high-attachment biases (as suggested above), it is not possible to take these results as clear evidence against this hypothesis. (See also Bradley, Fernández and Taylor (2003) for English data contradicting this hypothesis.)

6.4 Conclusions

The goal of the listening study was to assess the extent to which attachment decisions are affected by (overt) prosodic phrasing. This was achieved through examining the perception of overt prosodic differences in the 6 conditions created by a partial combination of three experimental factors: prosodic phrasing (Phrasing 0, Phrasing 1 and Phrasing 2), construction type (FS and CS) and RC length (short and long RC).

For the FS conditions, the data clearly show that prosodic phrasing has a dramatic effect on attachment interpretation. The three different prosodic patterns tested lead to very different attachment preferences for the same sentence: An FS sentence containing a short RC was interpreted with low attachment of the RC when the sentence had a prosodic break after N1 (Phrasing 1) and with higher attachment of the RC when the same sentence was read with a break after N2 (Phrasing 2). When no prosodic break was

present in the complex nominal with RC (Phrasing 0), there was a mild bias toward attaching low, which was presumed here to be due either to the absence of prosodic break, or to a syntactic Recency preference which becomes detectable when the prosody is neutral.

It should be noted that although FS sentences showed a prosodic effect, when looking specifically at the interpretation of Phrasing 2, the high attachment rate was 55.1%, which is a rather low attachment rate for an RC preceded by a prosodic break (see also Fernández, 2007, on English). Possibly, this is due to the fact that the Hebrew data are only for a short RC (1 prosodic word). There are listening data from other languages, using materials with longer RCs, in which a pre-RC break yielded a higher rate of high attachment interpretation: Stoyneshka (2007) found that Bulgarian listeners interpret a break before a long RC with a remarkably high rate of high attachment (82% high attachment), compared with a very low high attachment rate of 11% for a prosodic break between N1 and N2; Lovrić (2003) found that in Croatian, breaks before an RC (2 prosodic words in length) are interpreted by listeners as attaching 70.4% high on average. The Hebrew listening study did compare short and long RCs for the CS, but the results may not be extrapolatable to FS, see discussion below.[129]

The CS cases with a short RC further confirm the effect of prosodic phrasing on preferred RC-attachment, with Phrasing 2 more likely to result in a high attachment interpretation compared with Phrasing 0. However, although both constructions are affected in a similar direction by the changes in prosodic patterns (both FS and CS

[129] Another possibility is that the incidence of pre-RC breaks was overestimated in the production experiment either because the pre-RC position is the locus of the joint between the two contributing phrases or because the lack of fillers in the production experiment may have fostered routinized productions fitting a repeated "template".

sentences are interpreted higher with Phrasing 2, and lower with Phrasing 0), the difference in likelihood of high attachment is twice as large for the FS as for the CS. This is further supported by the finding that overall, the CS is interpreted higher than the FS in the listening experiment. In particular, for each of the prosodic phrasings (Phrasing 0 and Phrasing 2), CS sentences were interpreted higher than FS sentences (contrary to the prediction that similar prosodic phrasings would result in similar attachment preferences). These differences between FS and CS comparable sentences imply that internal properties of the CS tend to drive a high attachment interpretation of an associated RC, regardless of prosodic phrasing manipulations. Finally, the CS presents an unexpected high attachment preference, higher than 50%, for Phrasing 0, in contrast to the prediction that the locality of the RC, in the absence of a prosodic break intervening between N2 and the RC, would lead to a moderately low attachment interpretation of the RC, as observed for FS sentences produced with phrasing 0. (An alternative explanation would be that the low attachment rate in the FS is the unexpected finding (contra Recency), while the CS data for Phrasing 0 is expected; however, of the two choices, it seems that the more likely diverging data point is the one with the CS.) These findings conspire to suggest that there are other properties internal to the CS which encourage the RC to attach high, regardless of prosodic phrasing. One possibility is that the informational structure of the CS makes it more susceptive to high attachment modification: if the CS is processed as one informational unit, then N2 cannot be singled out informationally and therefore RC modification of N2, though syntactically possible, is less likely (see Chapter 8, §8.3 for a more detailed discussion of this and other avenues of thought which were considered in order to explain the CS data).

Other aspects of the data, in particular the manipulations of RC length, are revealing with respect to the contribution of prosody in processing. The CS data show that the likelihood of high attachment is strikingly similar, whether the RC is short (containing 1 prosodic word) or long (containing 3 prosodic words). The listening experiment found that short and long RCs are interpreted similarly when preceded by a prosodic break. As was noted above, this conclusion about the lack of length effect when a similar prosody is used was drawn based on CS data and so should be treated with considerable caution, since there are some grounds for thinking that CS is processed differently, and with less role for prosody, than FS. Therefore the possibility remains that, contrary to what seemed a reasonable assumption in planning the listening experiment (see §6.2.1) and despite my intuitions, the outcome for FS would have been different (but see Fernández' (2007) listening study of the RC construction in English, which found no length effect for pre-RC breaks).

To summarize: The listening study has provided data in support of the central role that overt prosody plays in determining attachment preferences for the ambiguous RC construction in Hebrew. It showed that different prosodic phrasing leads to different attachment biases for an identical sentence: the effect of prosody was overwhelming in the case of the FS, and was present, although somewhat reduced in the CS cases, suggesting that other, non-prosodic factors interact with prosody in determining attachment choices in the CS. The listening study confirmed that pre-RC breaks, associated in the production study predominantly with high attachment meaning and with long RC sentences, are interpreted higher than other phrasings, which were associated (almost) only with low attachment meanings in the production study.

To the extent that, as documented here so far, there are length and attachment interpretation effects in production of unambiguous sentences, and prosodic effects in perception of overt prosody in ambiguous RC sentences, it is of interest to investigate whether similar effects are found in a reading task, where any prosody that is present is projected by the reader (unless coded by punctuation). In silent reading, in particular, only implicit prosody could be at work. Chapter 7 presents data on the interpretation of the ambiguous RC construction in silent reading in Hebrew.

Appendix A to Chapter 6: Consideration of position for additional prosodic breaks prior to the complex nominal with RC.

The listening experiment aimed to investigate instantiations of the three phrasing categories found in the production study; these were hypothesized to result in varying interpretations for the ambiguous RC construction. However, the need to maintain the natural prosody of the target sentences as a whole required consideration of the prosodic structure of the whole sentence, i.e. including the material preceding the complex NP with RC sequence (i.e., the matrix Subject-Verb sequence). The position of an additional prosodic break (if any) prior to the complex nominal was considered carefully, based on findings from the phrasing data of the production study [130]. In what follows I consider for each phrasing category (*N1 N2 RC*, *N1)(N2 RC* and *N1 N2)(RC*) the particular instantiations of it that were found in the production experiment, and I argue for the choice of one of the competitors as optimal for use in the listening experiment. In particular, I argue in favor of the three phrasing patterns presented in (1) in the main text, and repeated in (1) below.[131]

(1) a. Phrasing 0: (S V) (N1 N2 RC)

b. Phrasing 1: (S V N1) (N2 RC)

c. Phrasing 2: (S V) (N1 N2) (RC)

[130] See Chapter 5 for a summary of the statistical analysis of the ear judgment data. For the percentage of use of each phrasing pattern observed in the production data, refer to appendix B to Chapter 5.
[131] The examples do not specify the length of the RC, which is addressed separately in Appendix 2 to this chapter. In this appendix, I discuss the consequences of each phrasing pattern for the minimum length (1 prosodic words), with any additional material within the RC marked by the plus sign (+).

Consider the phrasing category *N1 N2 RC*, as instantiated by Phrasing 0 in (1a). By definition, this category has no prosodic break within the whole sequence beginning with N1. The specific realization that was selected for use in the listening experiment had a single break following the main verb. Two possible phrasing patterns compatible with the definition of the phrasing category *N1 N2 RC* were observed in the production data, as shown in (2). (There were no utterances in the production study which contained no prosodic break at all.)

(2) a. (S V) (N1 N2 RC) 2-3+

b. (S) (V N1 N2 RC) 1-4+

In pattern (2a), the single prosodic break in the sentence follows the verb, thus creating a prosodic structure with minimally two prosodic words followed by three prosodic words, and a longer second phrase, if RC is longer than one prosodic word (2-3+); in pattern (2b) the break is after the subject, creating a pattern with one word followed by four or more prosodic words, depending on the length of the RC (1-4+). Clearly, pattern (2a) has a more balanced prosodic structure. The production study offers frequency evidence in support of the choice of this pattern: Among the *N1 N2 RC* phrasing category, a break after V is more common than a break after the subject in the low attachment data (overall 15.73% and 2.9%, respectively), and it is the only observed phrasing pattern in the high attachment data (3.1% overall). It should be noted, however, that the incidence of the phrasing category *N1 N2 RC* is low in general: it is observed practically only in short RC (30.4% in short RC overall; 1% in long RC overall). It is also more evident in the CS (20% overall) than in the FS (11.5% overall), presumably since the FS has another

available phrasing pattern (in which there is a prosodic break after N1), which was not observed in the CS data.

The phrasing category *N1)(N2 RC* by definition has a break after N1; the instantiation of this category which was chosen for the listening experiment was Phrasing 1 in (1b), in which a single break follows N1, and no other breaks occur in the sentence. Two main phrasing patterns compatible with this phrasing category were exhibited in the production data are considered as candidates for representing this category in the listening study, as presented in (3).[132]

(3) a. (S V) (N1) (N2 RC) 2-1-2+

b. (S V N1) (N2 RC) 3-2+

The pattern in (3a) has two prosodic breaks in the sentence: one after the verb and the other after N1, while the pattern in (3b) has only one prosodic break, after N1. These patterns were not observed in high attachment interpretation cases. The production data show that these patterns were only used in cases with forced low attachment interpretation. Frequency evidence reveals that both patterns have a small and roughly equal occurrence over all conditions (6.8% and 5.5%). These patterns were very infrequent overall: they were found to be common only with a short RC (the pattern in (3a) occurred in 12.5% and that in (3b) in 8.9%, of short RC utterances), and were rarely observed in the CS condition, presumably as consequence of the CS constituting a single prosodic unit: Short RC sentences were uttered with pattern (3a) in 4.2% of CS cases (cf.

[132] Other phrasing patterns were observed, as shown in appendix B to Chapter 5, but they were used in only a small percentage of utterances; shown here are the two most frequently used patterns observed in the production experiment.

20.8% of FS cases), and with pattern (3b) in only 3.2% of the cases (cf. 14.6% of FS cases).

The frequency of both phrasing patterns is quite similar, as evidenced by the production study. Thus it seems that the choice between the two instantiations of the *N1)(N2 RC* phrasing category in (3) for use in the listening experiment is not of great consequence. Intuitively, though, it seems more natural not to place a prosodic break both after the verb and after N1, which creates a 2-1-2+ pattern.[133] Instead, the prosodic phrasing which was chosed had a single prosodic break after N1, resulting in a 3-2+ pattern. Note that both (3a) and (3b) group the sentence into phrases that are not optimal according to the constraint ranking proposed in Shaked (2007) for Hebrew (see also Chapter 5, §5.4): (3a) violates *BinMin* and (3b) violating *BinMax*; Since it is argued that *BinMin* and *BinMax* are not ranked with respect to each other, it does not seem that the choice between the two phrasing patterns in 3 is consequential on these grounds either.

Finally, we turn to phrasing category *N1 N2)(RC*, which by definition has a break after N1. The specific realization that was chosen to be use in the listening study is Phrasing 2, as in (1c). In almost all instantiations compatible with the definition of this phrasing category, an additional break occurred prior to the N2) break.[134] The additional break either occurred after the verb, as shown in (4a) or after the subject, as in (4b).

---

[133] The prosodic structure in (3a) was found to be at least as common in production as the prosodic phrasing in (3b). It seems that the desire to convey a low attachment interpretation of the sentence by means of a break after N1 overcomes the requirement for prosodic phrases that are optimally 2 prosodic words long, even when no such break is predicted by the syntactic structure (see Chapter 5, §5.1).

[134] One additional phrasing, with no break at all prior to the N2] boundary was exhibited in the production study, but it represents only 0.025 of utterances. The only other variation was within the *N1 N2) (RC* phrasing category in the form of RC-internal breaks; these represent only 0.1% of long RC items in the production study. Since these phrasing were used so infrequently, they were not considered as phrasing options for the listening study.

(4) a. (S V) (N1 N2) (RC) 2-2-1+

b. (S) (V N1 N2) (RC) 1-3-1+

The production experiment offers frequency evidence in support of the phrasing pattern in (4a): Across all conditions (attachment height, RC length and construction type), pattern (4a) was the most common. In the forced-high attachment productions, pattern (4a) is used overall in 89.6% of the cases (85.4% and 91.7% for CS with a short and long RC, respectively; 86.5% and 94.7% for FS with short and long RC, respectively). In the forced-low attachment data, this pattern was used overall in 58.7% of the utterances (41.1% and 86.5% for CS with a short and long RC, respectively; 29.2% and 78.1% for FS with short and long RC, respectively). Even though it is less often manifested in the low attachment data, pattern (4a) is more common there than pattern (4b): Pattern (4b) occurred in only 5.7% of the production data overall. Thus, the frequency evidence supports the choice of the more common pattern (4a) for use in the listening experiment. Note that this pattern was found in both short and long RC productions, and therefore there is reason to believe that it is a natural prosodic structure in Hebrew. The discussion in Appendix B to this chapter of the choice of RC length for the listening study will draw support from these findings concerning placement of prosodic boundaries.

Appendix B to Chapter 6: Considerations in the choice of RC length

The following are arguments for the choice of an RC of 1 prosodic word for the listening experiment, in all cases except the explicit RC-length contrast for CS items with Phrasing 2. The possible RC lengths it is practical to test are 1, 2 or 3 prosodic words. The goal was to compare different locations of prosodic boundaries for an RC with a fixed length. Thus the practical question was what length of RC would be most neutral in order to best reveal general effects due to prosodic boundaries. The three different phrasing patterns must be considered separately, to assess the RC length that is most neutral for them.

For Phrasing 1, with a break after N1, the choice between an RC of 1 prosodic word versus 2 prosodic words does not seem to affect prosodic naturalness; both choices of RC length, as outlined in (1a) and (1b), seem equally natural when there is a break after N1.

(1)	a.	(S V N1) (N2 RC_1 RC_2)	3-3
	b.	(S V N1) (N2 RC_1)	3-2

Although the production study does not include data for an RC of two prosodic words, it may be useful to compare production data for short (1 prosodic word) and long (3 prosodic words) RC, in order to arrive at the most suitable RC length for the listening experiment. The production study had the aim of capturing natural prosody and therefore utterances collected in that study represent phrasing patterns typical for native Hebrew speakers. Basing the choice of phrasing patterns for the listening experiment on

tendencies found in the production experiment ensures to the best of our ability a natural sounding set of sentences for the listening experiment. The data from the production study show that a break after N1 occurs with both short (1 prosodic word) and long (3 prosodic words) RCs, but it is more common with a short RC (14.6% in FS-Short RC) than with a long RC (4.2% in FS-long RC). Data is cited for FS only, since the patterns in (1) are extremely uncommon in the CS, due to its status as a single prosodic word. Thus, use of a short RC is preferred for Phrasing 1 in the listening experiment. It should be noted that the production data lead us to hypothesize that Phrasing 1 will result in a low attachment interpretation regardless of the RC length (since it is almost only in the low-attachment conditions, both long and short FS, that this pattern was observed); therefore the choice of RC length is not likely to be crucial for Phrasing 1.

For Phrasing 2, with a break after N2, the choice of 1 prosodic word for the RC, as in (2b), may initially seem problematic, since phrasing 2 results in the RC constituting a prosodic unit on its own, creating what is intuitively a non-preferred prosodic pattern for a short RC. The perfectly balanced prosodic structure offered by an RC of two prosodic words, as shown in (2a), may seem preferable.

(2) a. (S V) (N1 N2) (RC_1 RC_2) 2-2-2

b. (S V) (N1 N2) (RC_1) 2-2-1

However, data from the production study show that a break after N2 is a highly common phrasing for RCs, even short RCs, in the high attachment condition; almost all the utterances (both FS and CS, with both short and long RCs) were produced spontaneously using Phrasing 2 in order to express the high attachment interpretation. Therefore, we

conclude that an isolated RC of 1 prosodic word, as in (2b), is a natural prosody for such sentences in Hebrew (though possibly not in other languages).

For Phrasing 0, with no breaks at all within the complex nominal plus RC, using an RC of 2 prosodic words as in (3a) is possible but it is less balanced than using a one-word RC as in (3b).

(3) a. (S V) (N1 N2 RC_1 RC_2) 2-4

b. (S V) (N1 N2 RC_1) 2-3

Intuition suggests that in normal use, a 2-4 prosodic structure as in (3a) is likely to be broken into a more balanced 2-2-2 structure if possible, in which case (3a) could sound unnatural. Therefore, it seemed best to use a single word RC for Phrasing 0.

After considering the different options for the weight of the RC, based on findings from the production experiment data, it was decided to use an RC containing a single prosodic word in the listening experiment. A single prosodic word is a natural RC length for all 3 phrasing patterns under investigation in this experiment. This choice is mainly based on the fact that a longer RC (of 2 or more prosodic words) is unnatural for Phrasing 0. Adding up the occurrences of the three phrasing patterns used in the listening experiment with a short RC shows that they cover more than 80% of the phrasings revealed in the production experiment, as Table 1 shows. (The phrasings that were omitted by and large represent variants of the prosodies studied, differing usually in the position of breaks prior to the construction of interest, e.g., a variant of Phrasing 2 with the pattern *(S) (V N1 N2) (RC)* rather than *(S V) (N1 N2) (RC)*. See Table 2 below for the full coverage of phrasing patterns observed in the production study.

Table 1: Percentage in the production experiment of occurrence and overall coverage of the three phrasing patterns (Phrasing 0, 1 and 2) chosen for the listening study (for short RC).

	Attachment	CS	FS
Phrasing 0 (S V) (N1 N2 RC)	Forced-low Forced-high	8.9% 5.3%	21.9% 4.2%
Phrasing 1 (S V N1) (N RC)	Forced-low Forced-high	N/A	14.6% 1.0%
Phrasing 2 (S V) (N1 N2) (RC)	Forced-low Forced-high	41.1% 86.3%	29.2% 86.5%
Coverage	Forced-low Forced-high Overall	80.0% 91.6% 85.8%	65.7% 91.7% 78.7%

Table 2: Coverage of phrasing patterns captured in the production study in each experimental condition, including the three phrasing patterns chosen for the listening study (in bold), and competing variant(s).

Category	Phrasing pattern	CS		FS	
		Low	High	Low	High
N1 N2 RC	**(S V) (N1 N2 RC)**	**38.9%**	**5.3%**	**21.9%**	**4.2%**
	(S) (V) (N1 N2 RC)	0%	0%	0%	1.0%
	Total:	38.9%	5.3%	21.9%	5.2%
N1)(N2 RC	**(S V N1) (N2 RC)**	**3.2%**	**0%**	**14.6%**	**1.0%**
	(S V) (N1) (N2 RC)	4.2%	1.1%	20.8%	0%
	(S) (V N1) (N2 RC)	0%	0%	5.2%	0%
	Total:	7.4%	1.1%	40.5%	1.0%
N1 N2)(RC	**(S V) (N1 N2) (RC)**	**41.1%**	**86.3%**	**29.2%**	**86.5%**
	(S) (V N1 N2) (RC)	6.3%	7.4%	3.1%	6.3%
	Total:	47.4%	93.7%	32.3%	92.8%
Coverage	**Phrasing 0,1,2 only**	**80.0%**	**91.6%**	**65.7%**	**91.7%**
	All variants	86.3%*	99.0%	94.7%*	99.0%

* Remaining portion of phrasing was found in phrasing pattern: (S) (V N1 N2 RC), which occurred in 6.3% of CS, and in 5.2% of FS utterances.

Appendix C to Chapter 6: Instrumental analysis of recorded stimuli

In order to examine instrumentally whether the recorded utterances used as stimuli in the listening study are natural and unexaggerated, it was necessary to establish a baseline comparison with spontaneously produced utterances of the same speaker. It was decided to use the Post-to-Times protocol (Bradley et al., 2003) in order to elicit utterances that would constitute that baseline.[135] Before the recording sessions of the stimuli for the listening study began, the yet-to-be-trained, naïve Hebrew speaker who was chosen for the recording of the stimuli for the listening study performed the Post-to-Time task used in the production study, thus producing spontaneous utterances corresponding to 8 of the 24 experimental items of the listening study. The trained speaker performed the task without interruption and produced fluent natural utterances. The procedural difference between her Post-to-Times utterances and the utterances she later recorded for the listening study was that, unlike the later recordings in which intended prosodic break positions were specifically marked, in the Post-to-Times recordings she was not instructed as to the placement of prosodic breaks. Recordings of these 8 items constitute the norm of natural utterances for the trained speaker and these were compared to her subsequent recordings of the materials for the listening experiment.

An instrumental analysis extracted duration means for N1 and N2 segments in the Post-to-Times recordings of the trained speaker, using the same procedure as the one used to analyze the recordings of the data from the production experiment. The

[135] The Post-to-Times protocol does not elicit completely spontaneous utterances, since it controls the content of the productions; however, it is the most effective method to my knowledge for producing natural-like semi-spontaneous utterances of a specific linguistic construction in a controlled setting, superior to merely having a speaker read aloud a pre-scripted sentence. In the absence of objective measures of naturalness, I submit with this dissertation (https://wfs.gc.cuny.edu/ashaked/www/index.html) the recordings from the production study reported in Chapter 5.

naturalness assessment first compared, for the 8 items tested with the Post-to-Times protocol, the duration means of the trained speaker's productions to the duration means obtained in the production study for 24 naive speakers, as reported in Chapter 5. This made it possible to determine whether the trained speaker's utterances resemble those of the participant population studied in this dissertation. A comparison between the trained speaker's duration means and duration means obtained in the production study is presented in Figures 1 for N1 durations and in Figure 2 for N2 durations.

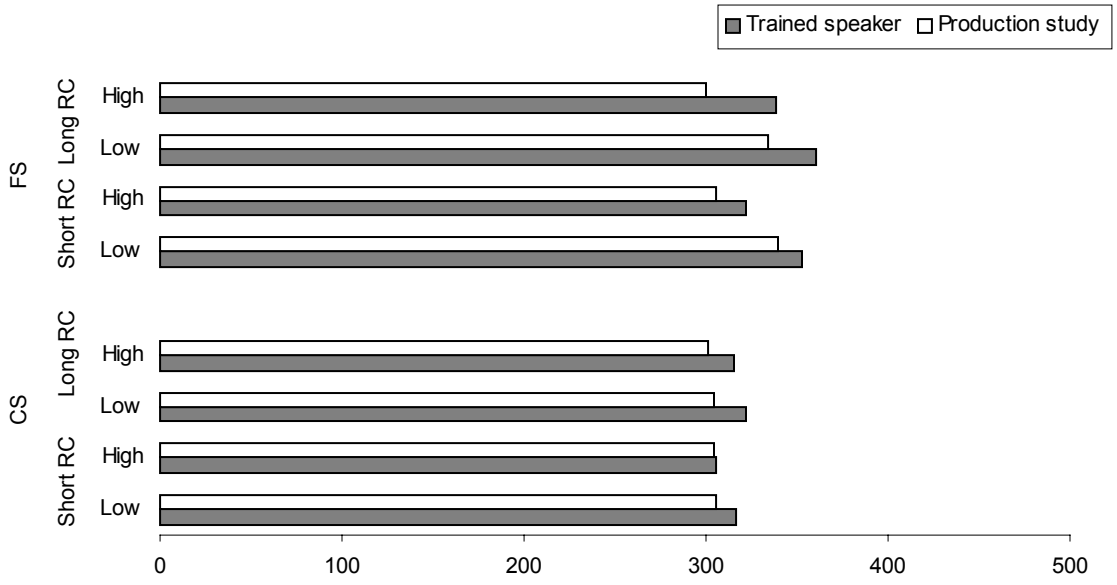


Figure 1: N1 Duration means (ms) for Post-to-Times utterances produced by the trained speaker, and those produced in the production experiment.

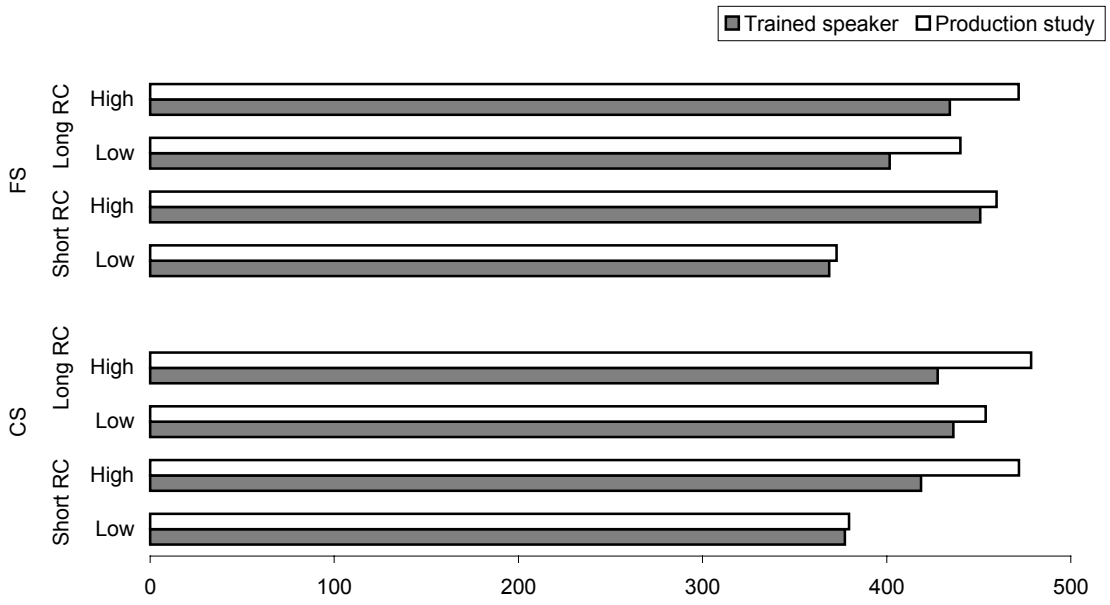


Figure 2: N2 Duration means (ms) for Post-to-Times utterances produced by the trained speaker, and those produced in the production experiment.

For N1 durations, the results show that the trained speaker's durations were generally slightly longer than the mean duration of N1 extracted from the comparable conditions in the production study reported in Chapter 5. The trained speaker's N1 durations were on average less than 15 ms longer than the means of the production study, with the exception of the FS long RC conditions, which showed durations that averaged 32 ms longer. Nonetheless, the mean durations of the trained speaker were not greater than 1 standard deviation above the mean durations in the production study. In addition, the result fell well within the limiting cutoff established in the analysis of the production study (Chapter 5, §5.2.5) to determine what constituted an acceptable lengthening with optional pausing (any utterance which was more than twice the mean duration of utterances in the production experiment was considered unnaturally long and hence was set down to the cutoff value in the production data analysis). For N2 duration, the means of the trained speaker were found to be somewhat shorter than the duration means in the

production study, in all conditions, although still within the normal range of durations observed in the production experiment. In sum, a comparison of the Post-to-Times productions between the trained speaker and the naïve participants recruited for the production study shows that the speaker's N1 and N2 durations fall within the acceptable range of durations produced in the production study. Thus it is possible to conclude that the prosody of the trained speaker, as reflected in N1 and N2 durations measured from her Post-to-Times recordings, is sufficiently similar to the prosody of the general population investigated in this study, and can be used as a measure for assessing the naturalness of the utterances she produced as stimuli for the listening study.

In order to assess whether the breaks produced in the recordings of the listening stimuli can be judged as natural, the Post-to-Times utterances were further analyzed. Perceived phrasing breaks were extracted for all Post-to-Times productions of the trained speaker, following the protocol used in the production study for extracting perceived phrasing judgments. Mean duration calculations were then made separately for each condition, in cases in which a phrasal break had been perceived to occur. This provided, for each condition, the duration means of those cases in which a break was perceived; these means were then compared with the duration means of stimuli utterances produced by the same trained speaker.[136] A comparison of N1 and N2 duration means between Post-to-Times utterances and listening stimuli utterances produced by the trained speaker, are presented in Figure 3 for N1 durations and in Figure 4 for N2 durations.

[136] The stimuli for the perception study were recorded with a specific prosodic phrasing, as the analysis in §6.2.2. shows. It was important to compare these utterances to comparable utterances from the Post-to-Times production, and therefore only cases where a prosodic break was perceived were taken into consideration.

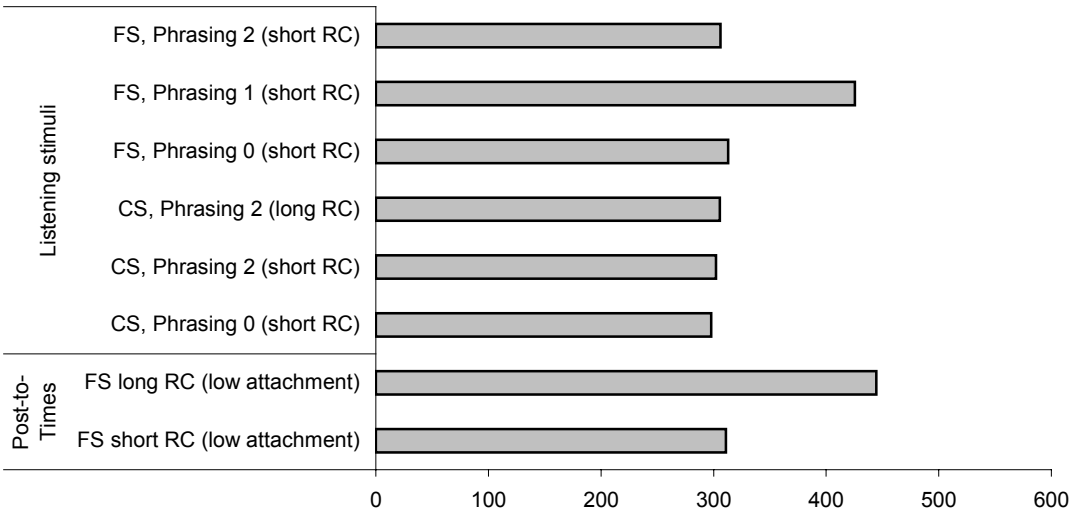


Figure 3: N1 duration means (ms) of listening stimuli in 6 conditions, and of Post-to-Times utterances in which a break was perceived, produced by the trained speaker.

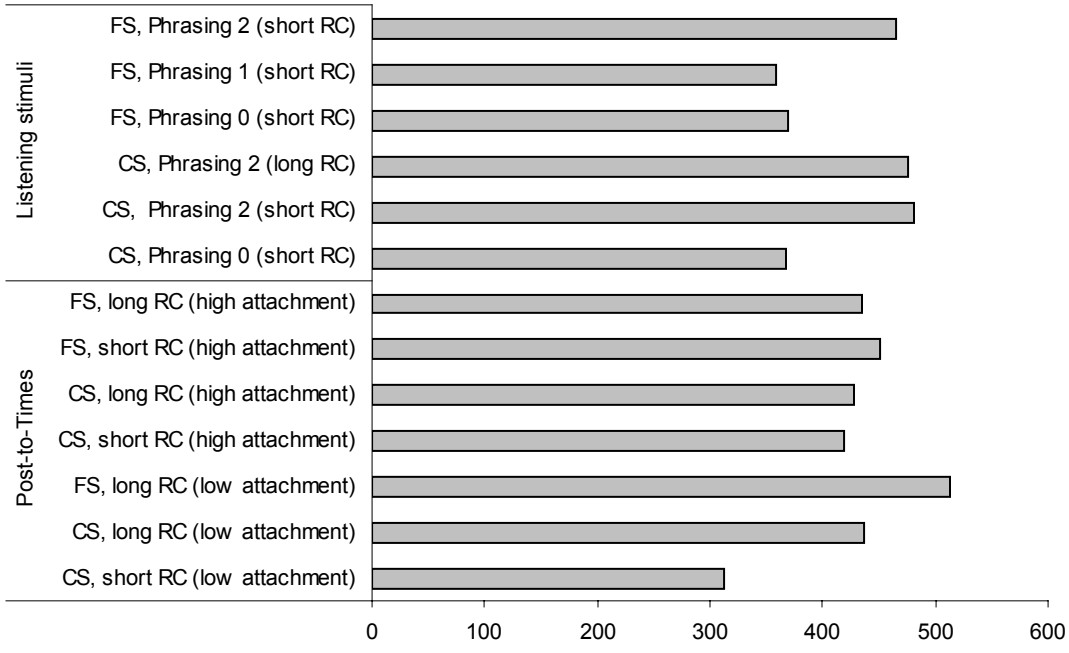


Figure 4: N2 duration means (ms) of listening stimuli in 6 conditions, and of Post-to-Times utterances in which a break was perceived, produced by the trained speaker.

A comparison was made of the means of N1 and N2 durations between the trained speaker's Post-to-Times utterances and the 8 equivalent items that she produced for the listening experiment stimuli. The comparison revealed that N1 and N2 duration were shorter for stimuli utterances than for comparable utterances in which a break was perceived in the Post-to-Time utterances produced by the same trained speaker; thus the results clearly show that the mean durations of N1 and N2 in the stimuli recordings fall within the average of the same speaker's natural utterances of Post-to-Times protocol in which a break was perceived. In sum, it is possible to conclude that the utterances used as stimuli in the listening study are natural, unexaggerated Hebrew utterances, placed comfortably within the range of utterances of naïve native speakers in the production study, and within the range of utterances produced by the trained speaker under the elicitation conditions of the Post-to-Times protocol; therefore there is no concern that they will distort outcomes of the listening study.

CHAPTER 7

ATTACHMENT IN SILENT READING

7.1 Aims

This dissertation focused thus far on the production and perception of overt prosody for Hebrew sentences containing the RC-attachment ambiguity. It is of value to relate overt prosody data to the interpretation of these ambiguous sentences in silent reading. Prosody has been claimed to have an effect on the interpretation of ambiguous sentences not only when it is explicit in the audible signal, but also when it is computed implicitly in silent reading (Fodor, 1998, 2002). The prosodic facts for the RC construction in Hebrew, as revealed by the production and listening data presented in Chapters 5 and 6, make interesting predictions for the interpretation of sentences read silently, on the assumption that a prosodic contour is computed implicitly. The silent reading experiment makes it possible to examine the relation between overt prosody and mentally projected prosody.

This chapter presents an attachment study in silent reading, which was designed to examine Hebrew speakers' ambiguity resolution preferences in attaching an RC to a preceding complex nominal. The attachment data gathered in this experiment reveal the range of attachment preferences for the two types of complex nominals (FS and CS) followed by RCs of two different lengths (short and long RC), and suggest implications about the phrasing in the mentally projected prosody, in relation to the overt phrasing

elicited in the production experiment (Chapter 5). It thus assesses the role that implicit prosody may play in affecting attachment choices for this construction in Hebrew.

The Implicit Prosody Hypothesis (IPH) proposed by Fodor (1998, 2002) states that, "In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction" (Fodor 2002:113). The implicit prosody assigned to a given construction is taken to be the same as its usual overt prosody, under the assumption that if a prosodic pattern is projected in silent reading, it is likely to be the same as the overt prosody that is most commonly employed by speakers for the same structure in an identical context. This cannot be tested directly, but is supported to the extent that it contributes to a satisfactory explanation of attachment behavior in reading. Given this background assumption, implicit prosody is claimed to be a cause of syntactic preferences in ambiguity resolution, and not just a consequence of independent syntactic or semantic attachment preferences.

The IPH predicts that the parser will prefer the syntactic structure that is associated (by the rules of the syntax/phonology interface) with the default prosody for that construction in that language. Thus, in order for the IPH to make any specific predictions about a given construction, the default prosody for that construction in that language needs to be determined empirically. For an unambiguous sentence, the default may be identified as whichever prosodic pattern is applied to it most consistently and often by speakers in a neutral context (broad focus). Identifying the default prosody for a syntactically ambiguous construction is theoretically and empirically more challenging.

An ambiguous sentence has two structures, and each structure may be associated with its own default prosody. It can be difficult to decide which of the two is the default of these defaults (i.e., which prosody is preferred before any choice of syntactic analysis has been made).

The production data reported in Chapter 5 reveal the possible prosodic structures associated with the different sentence types under investigation, which have been shown to be shaped by universal constraints on prosodic phrasing and their ranking in Hebrew (see Chapter 5, §5.4). These patterns in production may help identify the default prosody for the different experimental conditions tested.[137] The disambiguated elicited productions in the production study reveal that sentences with a long RC, both FS and CS constructions, pattern in much the same way, with a prosodic break preceding the long RC. This was true for sentences disambiguated for high and for low attachment, strongly implying that this pattern is the default prosody for long RC sentences following a complex nominal in Hebrew. For short RC sentences, in contrast, the prosodic phrasing patterns were quite different for high attachment and low attachment. For FS, all three major patterns (Phrasings 0, 1 and 2) were used with both high and low attachment though in very different proportions. For CS, Phrasing 1 was used only marginally, as expected in view of the CS' status as a phonological word; the remaining patterns

[137] Two aspects of prosody have been shown in previous work to affect ambiguity resolution. One is phrasing boundaries, which were shown (e.g., in Lovrić, 2003) to play a role in establishing the default prosody for this construction in Croatian. The other is focus, realized as pitch accent, which has also been shown to affect attachment interpretations in sentence comprehension for RC-attachment in English (Schafer et al., 1996, 2000). In attempting to establish the default prosodic structure for the Hebrew RC construction, I limit the following discussion to placement of prosodic phrase boundaries. Prosodic phrasing has been successful in offering explanations for German and Croatian. Although Schafer et al. (1996) provided evidence for the role of focal accent in English, Jun (2003) presented counter evidence which suggests that this is not the case, since a focused head noun did not attract RC attachment while "the existence of a prosodic boundary before a RC seems to be more important than the prominence in determining the attachment" (p. 235).

(Phrasing 0 and 2) were used, but in different proportions for different attachment interpretations. Therefore no consistent prosodic pattern could be identified as the default for the complex nominal construction per se, either for FS or for CS, in the short RC condition. The minimal use of Phrasing 1 in the CS does suggest that there will be more high attachment interpretations in CS than in FS, in general, since a break after N1 (Phrasing 1) is largely associated with low attachment interpretation, as discussed above.

The listening study presented in Chapter 6 provides data on the interpretation of overt prosodic patterns for the RC constructions, which show that manipulation of overt prosodic structure has a clear effect on the interpretation of otherwise similar FS sentences. In the case of the CS sentences, the effect was also present but was somewhat reduced, which suggests that prosody plays only a partial role in attachment interpretation in the CS. In order to test the IPH, the interpretation of ambiguous RC sentences in listening to spoken sentences is compared to ambiguity resolution in silent reading. This makes it possible to examine whether attachment preferences with overt prosody can be related to the preferences found in silent reading. A comparison between the data from the listening study (Chapter 6) and data collected in the silent reading study reported here (Chapter 7) will establish whether the overt and implicit prosodies are correlated. Even a strong correlation is no proof of a causal relation, of course, but it is at least a necessary condition on which a causal explanation might be built.

The IPH assumes that in the process of parsing sentences silently, readers first assign an implicit prosodic structure (based on prosodic rules particular to their language), and then interpret their own inner voice according to their inner ear. Thus, the process is viewed as sentence production (Chapter 5) followed by listening (Chapter 6),

inside one head. The production experiment (Chapter 5) showed that almost all long RCs had a break before the RC (Phrasing 2), and almost all high-attaching RC had a break before the RC (Phrasing 2). For the remaining condition, a short RC attaching low, the production data show equally likely occurrence of Phrasings 0,1,2 for FS and Phrasings 0 and 2 for CS (the near absence of Phrasing 1 in the CS productions was expected considering its status as a phonological word, as discussed in Chapter 2 and in subsequent chapters).[138] The listening experiment (Chapter 6) found an effect of prosody for FS sentences (only short RCs were tested). For CS, the listening data revealed a reduced prosodic effect for short CS sentences, with a surprisingly high attachment rate for sentences with Phrasing 0, and no length effect for Phrasing 2 sentences.

Given the production and listening data, it is possible to make the following predictions concerning the projected prosody in silent reading and its effect on interpretation for the two constructions under investigation. Consider first the sentences containing the FS. The standard sign of a prosodic effect in silent reading experiments is a length effect on the attached constituent: shorter (lighter) constituents have a tendency to attach lower than longer (heavier) ones (see Chapter 4 for references). The prosodic explanation relies on the assumption that a long RC, heavy enough to constitute a prosodic phrase on its own, would tend to induce a prosodic break before it, as confirmed in the production data. Then, the Structural Interpretation of Prosody Principle (SIPP)

[138] Originally, it had been expected that attachment differences between CS and FS would emerge solely as a result of their prosodic structure, with the CS constituting one phonological word and the FS two separate phonological words. This distinction did, as predicted, have a strong effect on the range of phrasing patterns produced, virtually eliminating the pattern with a break after N1 for the CS. Since a break after N1 yields very low attachment for FS in listening, it can be anticipated that the absence of this pattern in the CS will push up the relative proportion of high attachment in the CS compared with the FS, for a short RC. Thus, there is reason to expecting higher attachment for CS than for FS in the silent reading experiment. However, as discussed below, other factors seem to also be at play in determining attachment choice in the CS.

assumes "the existence of a skewing factor favoring a configurational interpretation of prosodic contours" (Fodor, 2002:129) and therefore would predict that the pre-RC prosodic boundary will be perceived by the parser as aligning with the right edge of a syntactic maximal projection which will favor high attachment of the RC to the whole complex NP. The SIPP does not predict 100% high attachment for pre-RC breaks, but only a bias towards high attachment (significantly more than 50%, assuming that materials are equi-biased). The IPH implies that this bias should be at work in silent reading (as in the arithmetics below, it is multiplied by the number of pre-RC breaks in silent reading, which can only sometimes be estimated from listening).[139] This explanation for the silent reading results in other languages suggests that Hebrew will behave similarly, with respect to the FS. Because we have seen that CS is less sensitive to prosody in listening, it will need to be discussed separately below (section 7.6). What the exact attachment preferences should be cannot be anticipated, for reasons that will be explained.

7.2 Method

7.2.1 Materials and design

The target sentences for the silent reading attachment study were the 24 experimental sentences which were used in the listening experiment (which included the

[139] SIPP is not an essential part of the IPH, in the sense that SIPP could be false even if indeed implicit prosody is projected in silent reading as the IPH claims. However, SIPP is essential to the IPH-based explanation of cross-language differences in RC attachment. The working assumption of the IPH is that in silent reading, each prosodic boundary projected by the reader is triggered by non-configurational factors such as length considerations, or category alignment constraints (e.g., $Align_L$ RC). If the reader always interpreted them as such, then they would not influence the structural configuration assigned to the sentence by the reader.

8 items used in the production study). The full list of experimental items is presented in Appendix A-4. The materials design for the silent reading study manipulated the two design factors: Construction type (FS and CS) and RC length (short and long RC), which were fully crossed. As in all the materials used for this dissertation study, long RCs were extended versions of short RCs, with additional material, such as time and place adjuncts, adding little information to the RC. Each of the 24 basic experimental items was presented in the 4 conditions created by the two factors. Example (1) demonstrates the 4 experimental conditions; (1a) shows FS with a short RC (1a), (1b) shows FS with a long RC, (1c) shows CS with a short RC and (1d) shows CS with a long RC.

(1) a. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref še-paraš.
the-fans admired ACC the-coach of the-wrestler who-retired

b. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref še-paraš
the-fans admired ACC the-coach of the-wrestler who-retired

le'axar ha-taxarut.
after the-fight

c. ha-'ohadim he'ericu 'et me'amen ha-mit'agref še-paraš.
the-fans admired ACC coach the-wrestler who-retired

d. ha-'ohadim he'ericu 'et me'amen ha-mit'agref še-paraš le'axar
the-fans admired ACC coach the-wrestler who-retired after

ha-taxarut.
the-fight

The 96 items (24 basic items x 4 experimental conditions) constituting the full set of target materials for the experiment were distributed across four questionnaire lists. Each list of the questionnaire took a different subset of six exemplars for each of the design's four item types, to avoid within-list repetition of lexical content. In addition to

the 24 target sentences, the experiment included 72 filler items. The filler items were similar to those used in the listening experiment, some exhibiting complex nominals and some containing RCs of different lengths (see Appendix A-5 for different types of fillers used), but importantly, they were all unambiguous. Similarly to the listening study, the silent reading attachment study had filler items in a ratio of 1:3 for experimental to filler sentences. Filler materials were identical across lists. They were used to disguise the purpose of the experiment, as well as to identify and exclude from the final analysis inattentive participants, whose responses on filler items was inaccurate more than 15% of the time.

An effort was made to maintain similarity between the fillers used in the listening study and those used in the silent reading study; however, minor changes were required by the design of the silent reading study. Filler items that had an unambiguously attached RC (24 of the 76 filler items) had a biased ratio of 1:5 short to long RC in the listening study, in order to match the 5:1 ratio of short to long RC in the experimental items design of the perception study.[140] In the silent reading study, since the experimental items had an even number of short and long RCs, the filler items were designed to mirror the even number of short and long RCs. Therefore, some of the long RCs of filler items from the listening experiment were shortened in order to allow for a ratio of 1:1 short to long RCs. This created a balanced number of short and long RCs in the filler and experimental

[140] The listening study had 6 experimental conditions, 5 of which had a short RC and only 1 had a long RC. To balance that bias towards a long RC in the experimental items, filler items with RC (which were only 1 type of several unambiguous filler items) had the opposite ratio of 5 long to 1 short RC, thus creating an overall equal number of short and long RC (see Chapter 6, §6.2.1). In the silent reading experiment, experimental items had an equal number of short and long RCs, therefore filler items also had a 1:1 short to long ratio.

items combined, in the silent reading study.[141] Other minor modifications were also made to about 1/3 of the filler items from their original form (without significant change to their content). A list of the filler items after revisions is presented in Appendix A-5.

The materials were arranged in a fixed pseudo-randomization, which was identical across lists in order to guarantee only minimal changes in the context in which a target sentence was presented. Thus, the experimental items did not change their position from one list to another. However, in each list the sentence appeared in a different one of the 4 conditions (FS/CS; short/long RC). The sentences were divided into 6 blocks. Each block contained 16 sentences: 4 experimental sentences and 12 filler sentences. The experimental items for each block included the 4 variations created by the factorial design, and were distributed such that in each block there were 2 items with animate nouns and 2 items with inanimate nouns. A block contained 4 filler sentences of each major type (2 of each sub-type).

The pseudo-randomization ensured that there were no more than 2 consecutive sentences of the same type and no more than 2 consecutive experimental items. In addition, there were no more than 3 consecutive items whose expected judgments would be on the same side (left or right) of the page, and no more than 3 consecutive items whose expected judgments would alternate right-side and left-side answers. The questionnaire was designed so that even though the expected judgments for each experimental item changed across lists (the judgment for filler items remaining the same across lists), the restrictions above were maintained. In ordering the sentences, obvious

[141] The following filler items were modified from their form in the listening study to accommodate the length criterion: 27, 29, 30, 37, 39, 40, 41 and 42. See Appendix A-5 for the full list of fillers, where modifications are noted for each filler item.

discourse connections (sentences involving a similar "scene") in adjacent or nearby positions were avoided.

A page-separator placed towards the middle of the questionnaire offered participants a short break, thus dividing the experiment into two parts. The pseudo-randomization ensured that the two parts of the questionnaire, following and preceding the optional break, started and ended with 3 filler items. In this way participants did not encounter an experimental item as the first or last sentence of each part of the questionnaire. In addition, each list began and ended with 3 covert filler items, for the same reason. All lists included 3 practice items for participants to respond to following the instructions (see Appendix A-5 for the practice items used in the silent reading experiment). Practice items were identical across lists. Including the practice items, each questionnaire included 99 sentences: 3 practice items, 24 experimental items and 72 filler items.

7.2.2 Procedure

The questionnaire had the form of a booklet. Each page contained one item. The sentence appeared on a single line, in the center of the page, as shown below in Hebrew for the sentence in (1a).

האוהדים העריצו את המאמן של המתאגרף שפרש.

Participants were instructed to read the sentence silently, at a comfortable pace that allowed them to understand its meaning, and then to turn the page. On the other side of the page there was a comprehension question about the sentence, with two answer choices. For target sentences, the question was about the attachment of the RC and the two answer choices were N1 and N2. The question remained in the same form whether the RC was short or long. An example of the follow-up comprehension question and answer choices for all 4 sentence-conditions in (1) is presented in (2), followed by the form of presentation in Hebrew.

(2)

mi paraš?	ha-me'amen	ha-mit'agref
who retired?	the-coach	the-wrestler

מי פרש?

א. המאמן ב. המתאגרף

For filler items, the sentence was followed by a question about some aspect of the meaning of the sentence; one of the answers was correct, while the other was incorrect. The two answer choices to the question that followed the filler sentences were plausible choices in general, but only one was correct in the specific context presented in the filler sentence (for example, for the sentence 'The producer argued with the screenwriter who insisted on adding a scene' the question 'Who insisted?' had the two answer choices 'the producer' and 'the screenwriter'. Although in general both can 'insist' and therefore are potential answers, in this sentence scenario only one is a correct answer choice, 'the screenwriter'). In cases where a filler sentence included a complex nominal, the question

did not refer to either of the nouns in the complex sentences. For target and filler sentences alike, the answer choices were presented in their order of appearance in the sentence.

Participants were instructed to respond to the question by circling the answer that best fitted the sentence they had read; they were not allowed to look back at the sentence once they had seen the question. For the target sentences, their answer reflected their preference for attaching the RC to one or the other of the two nouns in the complex nominal. At no time during the experiment were participants asked to read aloud, thus ensuring that if their answer choice was affected by what was arguably a prosodic factor, it must have been silent prosody created spontaneously while reading the sentences in a natural fashion.

Participants were tested in groups or individually, in a room within the facilities of the Linguistics Department at Tel-Aviv University in Israel. The experiment began with instructions and practice. The first pages of the questionnaire included instructions with guidelines for completing the questionnaire. Appendix B-3 presents the instructions used in the silent reading experiment. After a short practice session, participants were encouraged to ask any questions they had regarding the procedure of the experiment. After participants indicated that they understood the task, they were asked to complete the rest of the booklet in silence and with no interruption, except for the short break they were offered in the middle of the experiment. Participants received no feedback for their answers. The questionnaire took approximately 35 minutes to complete, including instructions and practice.

7.2.3 Participants

60 participants completed the experiment successfully (15 participants in each version). They were students at Tel Aviv University in Israel and were naïve with respect to the goals of the experiment. A background questionnaire similar to that used in the two other experiments ensured that participants were native speakers of Hebrew, and that Hebrew was the language they used in their everyday life. See Appendix C for the background questionnaire. Participants were monetarily compensated for volunteering to participate in the experiment, as for the two other experiments. Five additional participants were excluded from the data analysis due to inaccuracy on filler items (15% errors or more).

7.2.4 Data measurement and assessment

Responses to target items were coded in terms of attachment height, as attachment to N1 (high attachment) or to N2 (low attachment). Values for each item were then cast in terms of proportion of responses indicating high attachment. The design of the experiment allowed comparison of the 4 sentence types for each item, since they were presented at the same position in 4 different versions of the questionnaire, and differed only in the manipulated factors under investigation. A two-way analysis of variance was conducted, with Complex NP type (FS vs. CS) x RC length (long vs. short) design, for participants and for items. (All ANOVAs for the silent reading experiment are provided in Appendix D-3.)

7.3 Analysis and results

Summary data for the silent reading experiment are presented in Figure 7-1.

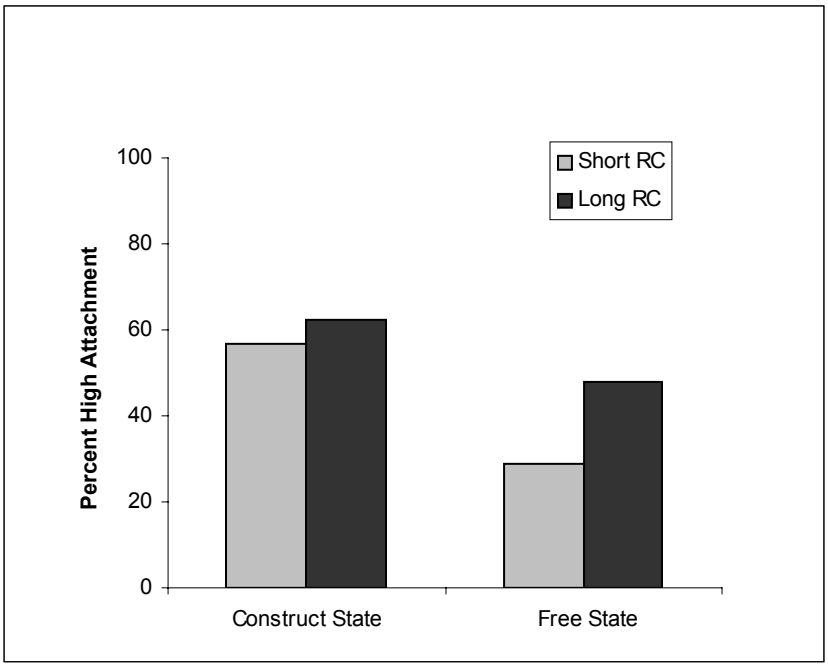


Figure 7-1: RC attachment to FS and CS in silent reading in Hebrew

Overall, attachment was higher in the CS than in the FS: participants interpreted the RC as attaching high 59.6% of the time with CS compared with 38.3% with FS ($F_1(1,56) = 46.13$, $p<.001$; $F_2(1,20) = 82.25$, $p<.001$). A main effect was also observed for RC length: participants attached high 42.6% of the time with short RC, and 55.3% with long RC ($F_1(1,56) = 37.89$, $p<.001$; $F_2(1,20) = 19.43$, $p<.001$). The main effects of construction type and RC length were found to interact ($F_1(1,56) = 11.97$, $p<.001$ and $F_2(1,20) = 5.43$, $p<.05$). Given a significant interaction term in the 2x2 analysis of variance, subanalyses of construction type were undertaken, looking separately at differences in attachment height between short and long RCs in the FS cases and in the CS cases.

In the FS, a significant RC length effect was found, with attachment higher for long RC than for short RC (F_1 (1,56) = 61.58, p<.001; F_2 (1,20) = 27.65, p<.001). The silent reading data thus show, as predicted, that readers interpret the long RC as having a greater likelihood to attach high (48.05% high attachment) than a short RC, while a short RC is interpreted as attaching fairly low (28.61% high attachment). It should be noted that although a long RC is interpreted higher than a short RC in Hebrew, it cannot be said that a long RC is interpreted high in absolute terms. The data show that a long RC is interpreted high roughly 50% of the time. This may seem low, but notice that it is in much the same range observed in the listening experiment. See detailed discussion below.

In the CS, on the other hand, there was no reliable length effect. Although sensitivity to length in the CS was in the usual direction (with long RC interpreted numerically higher relative to short RC), the RC length effect for CS was non-significant in the item-based analysis (F_2 (1,20) = 1.72, p>.10) and only a trend in the subject-based analysis (F_1 (1,56) = 3.39, .05<p<.10). Further examination of the data revealed that there were differences in magnitude of attachment height between FS and CS for each of the short and the long RC: the difference between FS and CS for short RCs is twice as large as the difference for long RCs (28 percentage points difference for short RC and 14.4 percentage points difference for long RC). Thus the significant interaction of the two main effects (reported above) shows that the sensitivity of length in the FS is notably weakened in the CS, suggesting that other factors, possibly inherent to the CS, override length manipulations in the CS (as already suggested in Chapter 6).

7.4 Summary of attachment findings

The results of the attachment experiment reveal the range of attachment preferences in silent reading of the ambiguous RC construction in Hebrew. The likelihood of a high attachment interpretation is greatest in the case of a long RC attaching to a CS (62.5%) and lowest in the case of a short RC attaching to an FS (28.6%). In earlier days, the classification of a language into a high attaching or a low attaching language was an important part of the general research program that maintained the universality of the processing mechanism in the face of apparent cross-linguistic differences. In recent years, however, the focus of discussion has shifted; it now centers less on cross-language differences and more on within-language variations in attachment preferences, which permits much closer control of lexical and other factors, increasing the precision of comparisons that may reveal influences of different factors on parsing decisions. If it were of interest nonetheless to reflect on the place of Hebrew in the high/low categorization of languages, then the construction used for the cross-linguistic comparison should be an FS with a long RC, since it is the most similar to constructions studied in other languages. With 48.05% high attachment this condition suggests that Hebrew in general is in the midline between high and low attaching languages. Still, it should be noted that a cross-language comparison can never be fully precise because the materials inevitably vary across languages, possibly presenting semantic, lexical frequency, or other biases; also, what constitutes a 'long RC' for purposes of the grammar may also differ from language to language.

The attachment study found a length effect for the FS cases, with attachment higher for a long RC and lower for a short RC, in line with the predictions outlined at the

beginning of this chapter based on the IPH's proposal that attachment in silent reading is prosodically influenced (as long as the long RC does not introduce semantic materials, the effect is taken to be because of prosody). For CS cases, however, the silent reading study did not find evidence of a length effect on attachment; although there was a numerical difference between short and long RCs corresponding to our expectations, it was not significant. The results for the CS suggest therefore that prosodic phrasing is not a major factor affecting attachment interpretation of CS sentences. The two constructions will therefore be discussed separately in what follows. Section 7.5 below integrates the outcomes of the three experiments for the FS, into a model that predicts the results of the silent reading experiment on the basis of the findings from the listening experiment and the production experiment. Section 7.6 presents the data for the CS construction obtained from the three experiments.

7.5 Integrating the outcomes of the three experiments for the FS

The silent reading data for the FS construction reveal an RC length effect: more high attachment interpretations were found for a long RC than for a short RC, as has also been observed in other languages (see references in Chapter 4). These findings are in line with the predictions following from an IPH account, as outlined in the beginning of this chapter. In silent reading, a long RC is predicted to be more likely to form a prosodic unit of its own and therefore induce a prosodic break preceding it as in the overt productions in the production study. This prosodic phrasing is characteristic of a high attachment syntactic configuration for the RC and so by SIPP it is predicted to result in a bias towards high attachment interpretation (significantly more than 50% high attachment).

Although the attachment data do show higher attachment interpretation of the long RC, some additional explanation is needed for the fact that even in that case the high attachment did not exceed 50%. This point is addressed later in this section.

The FS attachment data in silent reading can be accounted for in terms of implicit prosody, with a remarkably good fit, by relating them to the results of the experiments on overt prosody (production and listening). It is not possible to make exact numerical predictions, but a range of numerical combinations are considered below. It should be noted that in order to simplify the calculations, we take a strong version of the IPH. For convenience, the data from all three experiments are assembled together in Table 7-1 below (P0 stand for Phrasing 0, and so forth).

Table 7-1: Summary of results for FS production, listening and silent reading experiments. The measures here mix prosody rates (in the production study) and attachment rates in the listening and silent reading studies.

	Production						Listening			Silent reading
	Forced low			Forced high						
	P0	*P1*	*P2*	*P0*	*P1*	*P2*	*P0*	*P1*	*P2*	
Short RC	27.1	40.6	32.3	6.2	1.0	92.8	39.4	12.5	55.1	28.6
Long RC	1.0	15.6	83.3	0.0	0.0	100.0				48.1

Consider first the FS with a long RC. In production, this had a very consistent prosodic structure, with a break between N2 and the long RC, almost regardless of the intended attachment interpretation. This seems adequate justification for taking Phrasing 2 to be the default prosody for long RC with FS, which would be projected by the participants in the silent reading study. An exact percentage for the incidence of this pattern in silent reading cannot be established based on the production data, because the productions were differentiated by intended attachment height. Our working assumption

for the silent reading situation will be that no attachment choice has been made independently of prosody; it is the default prosody that will determine attachment. This very strong statement of the hypothesis is no doubt too extreme for real world data, in which there are many potential slippage points between the sentence form, the projected prosody, and then the preferred attachment; but in this section the research program is to pursue this strong version of the interface correlation in order to see how much of the data it can account for.

Long RC FS sentences show 100% of Phrasing 2 for forced high attachment, and 83.3% of Phrasing 2 for forced low attachment. Our working hypothesis entails that we should assume that in silent reading the percentage of Phrasing 2 projected by readers for long FS ranges between a lower bound of 83.3% and an upper bound of 100%. If it is valid to assume that a parser with no prior bias (i.e., no bias independent of prosodic preferences) toward either attachment would split the difference between these rates of Phrasing 2, the assessment would be 91.7% for long RC. Let us therefore compute anticipated rates of attachment in silent reading, on the basis of these three values of interest for Phrasing 2: 83.3%, 91.7%, 100%.

The listening data provide information on the attachment interpretation associated with the Phrasing 2 pattern. For the FS construction with Phrasing 2, the listening study found that 55.1% of the cases were interpreted as attaching high. Unfortunately, FS sentences with long RCs were not tested in that experiment. The only data available are for short RCs. One way to estimate the attachment rate of Phrasing 2 for long RC in listening is to assume that it would be similar to that of the observed short RC (55.1%); this working assumption supposes that with prosody established overtly, attachment

would be similar regardless of RC length, as was found for English (Fernández, 2007). Indeed, the CS data in listening show similar attachment height for short and long RCs with Phrasing 2. However, the CS listening data cannot be taken to provide absolute support for this assumption, since it was found that the CS is less prosodically sensitive than the FS, and is suggested to have an intrinsic bias towards high attachment (see Chapter 6, §4); this will be discussed further below. Nevertheless, the 55.1% value is worth considering as a candidate in the computations that follow, where it can be compared with other values. Of particular interest is a contrast with an alternative way of estimating the attachment rate for a long RC, which assumes, following from the Rational Speaker Hypothesis (RSH) (Clifton et al. 2006), a lower attachment rate for long than for short RC in listening.

The question of interest is whether the rate of high attachment for Phrasing 2 (e.g., 55.1%), together with any of the range (83.3% to 100%; mean 91.7%) of Phrasing 2 for FS long RC, predicts the silent reading attachment data.

X = 55.1% x 83.3% = 48.7%
X = 55.1% x 91.7% = 53.5%
X = 55.1% x 100% = 55.1%

The predicted percentages for 55.1% are 48.7%, 53.5%, 55.1%. The observed percentage for FS long RC in silent reading was 48.1%. Though on the low side, this seems to fit well enough with the range calculated based on the 55.1% estimate. If our theoretical interpretation is along the right lines, it also implies that the incidence of Phrasing 2 when attachment is not dictated by the input (i.e. in reading) is closer to that of forced low attachment than it is to that of forced high attachment – though these differ very little.

But in any case, the observed 48.1% high attachment for FS long RC in silent reading is within reach of the bounds calculated on the basis of IPH's explanation of attachment preferences in silent reading as due to the projection and interpretation of prosody comparable to the overt prosodic patterns in spoken sentences.

Another way of looking at the very same facts is to begin with the assumed percentage of high attachment for Phrasing 2 for long RC in the listening study (55.1%), together with the percentage of high attachment for FS in silent reading (48.1% high), and calculate from these the estimated number of times (X) that the Phrasing 2 pattern would have had to have been used in silent reading if the IPH explanation is correct.

$$55.1\%\ X = 48.1\%$$
$$X = 87.3\%$$

This estimate of 87.3% for pre-RC breaks in silent reading is to be compared with the observed breaks before RC in overt prosody in the production experiment: 100% for high attachment and 83.3% for low attachment. While clearly 87.3 is not identical with either number, it is within their range (and as above is closer to the forced low attachment case than to forced high attachment), and so this estimate may be regarded as supporting the IPH hypothesis that prosody mediates the relation between silent reading and attachment interpretation.

As noted, the above explanation concerning long RC was forced to base its estimate of percent high attachment for Phrasing 2 on the listening data for short RC. This was necessary because the listening experiment unfortunately did not test the FS construction with long RCs (for design reasons outlined in Chapter 6, §6.2 above). The observed rate of high attachment for pre-RC breaks in FS thus derives only from short

RC cases. There do exist attachment data for long RC with Phrasing 2 in the CS construction (69.2% high attachment). But this is a less relevant benchmark since attachment is systematically higher in the CS than in the FS where there are data for both. As noted above, the tentative assumption in the above calculations was that with prosody established overtly, RC length is irrelevant to attachment; specifically, a break before an RC has similar consequences in terms of attachment regardless of the length of the RC.[142] Although the listening data for CS indicate that short and long RCs preceded by a prosodic break have almost identical interpretation tendencies (69.4% and 69.2% high attachment, respectively), this cannot be taken as direct support for a comparable assumption for the FS, because there is reason to suspect, as discussed above and in Chapter 6 (§6.4), that these CS results are partly driven by an inherent bias towards high attachment in the CS. Nonetheless, as we have seen, the estimate of attachment for FS long RC in silent reading which assumes the same attachment consequences for similar prosodic patterns regardless of the length (prosodic weight) of the RC, does fit the experimental data quite well.[143]

It is possible to yield a good fit by trial and error experimentation with estimated rates of Phrasing 2 in silent reading as well. For example, if the rate of high attachment

---

[142] It should be noted that Lovrić (2003) found that with prosody established overtly, the +/-Prep factor manipulated in his listening experiment on Croatian was irrelevant to attachment preference. Lovrić interpreted this as evidence that the primary determinant of attachment is prosody. Similar reasoning applies in the present case (Lovrić did not vary RC length in his listening experiment).

[143] The silent reading data for the FS could also follow from a frequency-based (non-prosodic) model for silent reading, such as the Tuning Hypothesis (Brysbaert & Mitchell, 1996; Mitchell, Cuetos, Corley & Brysbaert, 1995). This hypothesis proposes that initial parsing decisions (between high and low attachment) for ambiguous constructions are exclusively based on the relative frequencies with which high and low attachments have been encountered in comparable structures in the past. Suppose corpus data show the FS sentences with a long RC attach higher than FS sentences with a short RC; then, the probabilistic model would predict that long RCs would be interepreted higher than short RCs, based solely on the relative frequency with which each RC length was interpreted as attaching high or low in the past. The prediction of higher attachment rates for long RC compared with short RCs is as observed in the silent reading experiment. (Note that this explanation does not relate the attachment behavior to any property of the language, or of the construction; see also discussion in Chapter 8, §8.3)

for Phrasing 2 were 50%, then the three estimates would be 41.7%, 45.9%, 50%. The latter, which assumes a 100% association of high attachment with Phrasing 2, comes closest to the observed value. Of course, all of these are extremely approximate estimates. But it seems that the assumed 55.1% attachment rate for long RCs in response to Phrasing 2 utterances, together with 83.3% of Phrasing 2 in silent reading, fits better; in other words, assuming a similar attachment rate for short and long RC in silent reading provides a better fit for these calculations.

Turning now to short RC: Is it possible to explain the silent reading attachment data for a short RC with an FS, in terms of implicit prosody using the simple model proposed above for long RC? In contrast to the long RC, the elicited utterances of a short RC exhibited different phrasing patterns for the different attachment interpretations. For forced high attachment there was a strong bias (92.8%) towards Phrasing 2; for forced low attachment the three phrasing patterns were fairly evenly balanced, with no strong bias toward any one. Thus, based on the production data, it is not possible to arrive at a clear judgment as to a default prosodic structure for the short RC condition. Even though Phrasing 2 is numerically most frequent for the short RC in the production experiment, it is conceivable that it is a somewhat unnatural prosody for a short RC, which speakers employ mostly when it is needed to disambiguate high RC-attachment, but which they would not spontaneously adopt more frequently than any other pattern in silent reading. The best that can be done, in this case, is to consider several possible mixtures of the three prosodies, and find out post hoc whether there is one (or more) which fits the silent reading data satisfactorily.

Let us first consider the proportions of the three prosodies that were observed for short RC overall (i.e., combining both forced attachments in the listening experiment. For this estimate, we take the average of forced low and forced high conditions. That is: Assume that (27.1+6.25%)/2 of sentences have Phrasing 0, which has an acceptance rate of 39.4%; and that (40.6+1%)/2 of sentences have Phrasing 1, which has an acceptance rate of 12.5%; and that (32.3+92.8%)/2 of the sentences have Phrasing 2, which has an acceptance rate of 55.1%. Then we arrive at an estimate of 43.7% high attachment interpretation in silent reading of short RC with FS. The observed value is 28.6%, so this estimate is distinctly too high.

Let us now consider the possibility that prosody projection in silent reading is completely unbiased toward one phrasing or another. This amounts to assuming that one third (33%) of sentences have Phrasing 0, which has an acceptance rate of 39.4%; and that 33% of sentences have Phrasing 1, which has an acceptance rate of 12.5%; and that 33% of the sentences have Phrasing 2, which has an acceptance rate of 55.1%. Then we arrive at an estimate of 35.3% high attachment interpretation in silent reading of short RC with FS. This estimate is still somewhat too high, compared with the observed 28.6%.

The way to lower it would be to increase the estimated proportion of Phrasing 1 or to decrease the estimated proportion of Phrasing 2. Both, but perhaps especially the latter, seem perfectly reasonable for a short RC with no forced attachment, but we have no way to select particular percentages without introducing arbitrariness. Let us simply try out some possibilities. Consider the proportions 40%, 40% and 20% for Phrasings 0, 1 and 2, respectively. This yields an estimate of 31.8% high attachment in silent reading, which is closer to the observed 28.6%. A still lower estimate results from assuming the

proportions 45%, 45% and 10% for Phrasings 0, 1 and 2 respectively. This yields an estimate of 24.4% high attachment in silent reading, which is now lower than the observed 28.6% for the first time.

To summarize: It is in the nature of the case that no specific prediction can be made for the attachment preference of short RC in silent reading. These numerical exercises can establish only that there are some plausible combinations of phrasing choices in silent reading which would yield approximately the observed proportion of high attachment for short RC, and which reflect the evidence from the production experiment, that short RC are assigned more of the low-attaching Phrasings 0 and 1 and less of the high-attaching Phrasing 2, than in the case of long RC. The short RC data therefore do not add strong support to the IPH-based account, but they are fully compatible with it. Together with the prediction for the long RC, discussed earlier, it is clear that the silent reading data for FS fit comfortably with the IPH approach.

7.6 Understanding the data for the CS

The silent reading attachment data for CS differ in two major ways from the corresponding data for FS: (a) the CS show significantly more high attachment than the FS (for both short and long RC); (b) the CS does not show a significant effect of RC length: short RC were attached numerically less high than long RCs but the difference was not reliable. Can these points be explained on the basis of results of the production and listening experiments? The data from all three experiments are assembled together in Table 7-2 below, for convenience (P0 stand for Phrasing 0, and so forth).

Table 7-2: summary of results for CS production, listening and silent reading experiments. The measures here mix prosody rates (in the production study) and attachment rates in the listening and silent reading studies.

	Production						Listening			Silent reading
	Forced low			Forced high						
	P0	*P1*	*P2*	*P0*	*P1*	*P2*	*P0*	*P1*	*P2*	
Short RC	45.2	7.4	47.4	3.1	1.0	95.9	61.4		69.4	56.7
Long RC	5.3	1.1	93.7	0.0	0.0	100.0			69.2	62.5

The production experiment confirmed the familiar claim that Phrasing 1 is unnatural for CS. For long RC with CS, Phrasing 2 was used almost exclusively; for short RC with CS, Phrasing 0 and Phrasing 2 were used more or less equally often. Since Phrasing 1 yields the most low attachment for FS (which is where it can legitimately be tested), it could be predicted that CS would exhibit less low attachment in silent reading than FS. So point (a) is understandable. Note that this consideration should apply especially to short RC since the FS exhibit a substantial number of Phrasing 1 only for short RC. So this approach also offers an explanation for why the difference in attachment preference for CS and FS is almost twice as large for short RC as for long RC.

However, it is also important to note that in the listening study, the overt prosodies Phrasing 0 and Phrasing 2 both received more high attachment interpretations for CS than those same overt phrasing patterns did with FS. This was unexpected, but it provides another reason for anticipating higher attachment for CS than FS in silent reading.[144] For this prediction it is not even necessary to speculate about what mixture of Phrasing 0 and Phrasing 2 the CS might be assigned in silent reading. Note the implications for IPH. The striking difference in attachment preference between CS and

[144] In fact, the data suggest a hypothesis according to which the CS has a fixed rate of approximately 60% high attachment, totally regardless of prosody, in both listening and reading data. This simple hypothesis is however ruled out by details in the data, e.g., in listening, a significant difference of 8 points was found between Phrasing 0 and Phrasing 2.

FS in silent reading is compatible with the IPH assumption that similar prosodic forces are at work in reading as in listening; this is not undermined at all by the unexpected outcome of the listening study. It is the latter that stands in need of some additional type of explanation, such as tentatively outlined in Chapter 6.

Now to point (b). The long RC version of CS would be expected to have predominantly Phrasing 2, based on the production data (approx 98% of utterances with Phrasing 2 across high and low forced attachment); clearly Phrasing 2 is the default. In the listening study, the long RC with CS and Phrasing 2 showed 69.2% high attachment. Together with 61.4% for Phrasing 0 (observed for CS with short RC) and 12.5% for Phrasing 1 (estimated from the FS data), both negligible, this predicts 68.9% high attachment for long RC in silent reading. In fact the long RC CS shows a rate of high attachment (62.5%) in silent reading which is not incompatible with this prediction: it is a little lower than the predicted 68.9% but not too distant from it.[145]

In sum: For long RC, comparison of the silent reading data for CS and the listening data for CS gives a fairly satisfactory fit. For the short RC it is not possible to make a clear prediction because it exhibits two different prosodic phrasings in the production study. The best that can be done is to compute some estimates and see whether any are compatible with the reading data, as we did for the FS above.

Let us first consider the proportions of the three prosodies that were observed across both forced attachments in the listening experiment. For this we compute the average of forced low and forced high conditions. That is: Assume that (45.2+5.3%)/2 of

[145] However, as for FS, the listening data for Phrasing 0 in the CS are available only for short RC and must be extrapolated to provide an estimate for long RC. Since there is clear evidence in the listening data for Phrasing 2 that the attachment preference in CS is unaffected by RC length, the extrapolation for Phrasing 0 is not unwarranted.

sentences have Phrasing 0, which has an acceptance rate of 61.4%; and that (8+2%)/2 of sentences have Phrasing 1, which has an acceptance rate of 12.5% (borrowed from FS); and that (47.4+93.7%)/2 of the sentences have Phrasing 2, which has an acceptance rate of 69.4%. Then we arrive at an estimate of 65% high attachment interpretation in silent reading of short RC with CS. The observed value is 56.6%, so the estimate is too high. Another possibility to examine is that prosody projection in silent reading is completely unbiased toward one phrasing or another. For CS, let us assume that Phrasing 1 is excluded, so this hypothesis amounts to one half (50%) of sentences having Phrasing 0, which has an acceptance rate of 61.4%; and 50% of sentences having Phrasing 2, which has an acceptance rate of 69.4%. Then we arrive at an estimate of 65.4% high attachment interpretation in silent reading of short RC with CS. This estimate is slightly higher than the previous one, hence too high compared with the observed 56.6%.

In fact, with Phrasing 1 excluded, it is impossible for the estimate to be any lower than 61.4% for silent reading (as with 100% Phrasing 0). This is because there are only two relevant prosodies and both yield high attachment rates over 60% in the listening study. It is inevitable, therefore, that the short RC with CS has somewhat lower attachment in silent reading than anticipated on the basis of overt prosody in production and listening. The mismatch is not very extreme however. One aspect of the silent reading data that is well captured by these estimates across experiments is that for CS the short RC has only a slightly fewer high attachment rate than the long RC. This is because the attachment rates in *listening* are so similar across RC lengths and across prosodic patterns. Thus the silent reading data can be assimilated; it is the high attachment rates for CS *in listening*, and their insensitivity to RC length, that remain puzzling.

To summarize, the current chapter on RC attachment in silent reading revealed length effects in the FS, and it was argued that the data presented in this chapter can be accounted for neatly by invoking the implicit prosody hypothesis. A similar length sensitivity was not observed for the CS in silent reading, but this was the result of the puzzling results of the CS in listening, which need further explaining. Chapter 8, §8.3 will explore several hypotheses, which might account for the unexpected and puzzling behavior of the CS.

CHAPTER 8

CONCLUSIONS

The research reported in this dissertation contributes to the study of how syntactic processing is influenced by prosody. There has been a burgeoning of studies on this point in the recent literature (see references in Chapter 6, §6.1). A topic of particular importance is whether the relationship between prosodic and syntactic structure varies from one language to another. Most especially, it is of interest whether any difference in the interface relationship can be traced directly to the grammars of different languages, or whether it is an arbitrary relation that varies unpredictably from one language to another. Ideally, the grammar-specific properties of the prosody-syntax interface will prove to be an aspect of universal grammar (UG), which language users do not need to acquire on the basis of experience. If it is a principled relationship, then that would give rise to a much more constrained theory of the nature of human language than if the syntax-prosody interface could vary freely from one language to another. For this reason it is extremely important to examine the syntax-prosody interface in a variety of languages, and especially to systematically study languages or constructions that apparently differ from each other in interface-relevant properties and can therefore shed new light on these issues.

The present study of the syntax-prosody interaction in complex nominals in Hebrew was undertaken because Hebrew has several types of complex nominals and thus offers a useful set of syntactic and phonological contrasts. Two of these – the Free State

(FS) nominal and the Construct State (CS) nominal – were used here to study in detail the relationship between attachment preferences in Hebrew and the role of prosodic phrasing in speech and in silent reading. This project adds to the results of previous ones in several respects: (1) It is the first study of sentence processing in Hebrew; (2) The Hebrew language provides a cross-construction contrast which is particularly suited for investigating prosodic effects on processing; (3) The research program links data from production, perception and silent reading experiments with carefully matched materials; (4) The findings of the production data are based on both a ToBI-like phrasing analysis and instrumental analysis of prosody (using both duration and f_0 data), which strongly confirm one another; and (5) The accumulating data identifies the CS as a unique construction in terms of its attachment behavior, in contrast to constructions in other languages studied thus far.

This concluding chapter recapitulates the central ideas presented in the preceding chapters, first by describing the motivation and aims for the investigation, and then by briefly restating the main results of the three experiments. The summary of experimental findings discusses each linguistic construction separately. I first present the data for the FS construction, since the FS is the most comparable to related constructions in other languages that have been studied and reported in the literature. I then move on to present and compare the data for the CS construction. This construction does occur in other languages (e.g., Standard Arabic) but it has not previously been subjected to psycholinguistic investigation. I go on to address the puzzling CS data, and consider several hypotheses that may explain the inherent bias towards high attachment in this complex nominal. Finally, I raise some open questions as well as possible routes along

which future research may proceed in the investigation of Hebrew prosody and its relation to sentence processing.

8.1 Motivation and aims

For many years now, great focus in the psycholinguistic literature has been put on the question of the mechanism underlying the processing of the ambiguous RC construction. In sentences containing this construction, one of the two nouns in the preceding complex nominal, i.e., either the whole complex nominal (whose head is N1) or the second nominal (N2) by itself, is modified by the RC. This is illustrated in (1) with a Hebrew FS complex nominal.

(1) ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref še-paraš
the-fans admired ACC the-coach of the-wrestler who-retired
'The fans admired the coach of the wrestler who retired'

Research on this construction in diverse languages has shown that while the interpretive preference in some languages is for modification of the whole complex nominal, headed by the first nominal in the sequence, in other languages it is preferred to interpret the RC as modifying the second noun on its own. As noted in Chapter 1, there has been a collection of explanations of cross-linguistic approaches to explaining attachment preferences, many of which have been disconfirmed by Lovrić (2003) and Augurzky (2005). See Lovrić (2003) for a summary of the different proposals. Among the proposals which remained in play, the Tuning hypothesis (Cuetos & Mitchell, 1988; Mitchell, 1994; Mitchell et al., 1995) was abandoned some years ago, though there has been some attempts since to reconcile the corpus data with the experimental data (e.g., see Desmet et

al., 2002, 2005). The Anaphoric Binding dualism proposal of Hemforth et al. (1998, 2000a, 2000b) was falsified in Augurzky (2005), and Hemforth and colleagues' (Hemforth et al., 2002, 2006) explanation based on informational structure was challenged by Fernández and colleagues (Bradley, Fernández & Taylor, 2003; Fernández, Bradley, Igoa & Teira, 2003; Fernández & Bradley, 2004).

In recent years a great deal of the research has focused on a prosodic explanation of the variation in attachment choices within and between languages. Since the original data were from silent reading experiments, this line of explanation began with the proposal of the Implicit Prosody Hypothesis (IPH) (Fodor, 1998, 2002), that the attachment choices in silent reading are affected by the implicit prosodic contour that is computed for the sentence by the reader, on the basis of grammar-specific phonological rules in the language. Since implicit prosody cannot be directly observed, it could be investigated only by studying of the influence of overt prosody on RC-attachment in speaking and listening. There are now experimental data on overt prosody and RC processing in several languages (Croatian: Lovrić, 2003; English: Maynell, 1999; Bradley, Fernández & Lovrić, 2003; German: Augurzky, 2005; Japanese: Jun & Koike, 2003; Korean: Jun & Kim, 2004; Hwang & Schafer, 2009; and Spanish: Fernández et al., 2003). This dissertation provides data on prosodic phrasing in Hebrew and its relation to the processing of RC-attachment in two types of complex nominals which have different prosodic characteristics.

First, a brief re-cap of the main properties of the two complex nominals. They share the same meaning, though they differ in their syntactic and phonological properties. The FS is presented in (2a) and the CS is presented in (2b).[146]

(2) a. ha-me'amen šel ha-mit'agref
the-coach of the-wrestler

b. me'amen ha-mit'agref
coach the-wrestler

a. and b. 'the coach of the wrestler'

The FS is comparable to the prepositional genitive construction available in many languages in which the RC construction has been studied (e.g., English). It has two nominals linked by the preposition *šel* ('of'). Importantly, each of the nominals within the FS is a full DP and constitutes its own prosodic word. The CS, on the other hand, does not have a preposition linking the two nouns, and the head of the complex nominal appears in a morphologically bound form and is syntactically restricted in several ways: it cannot be marked separately for definiteness but rather inherits its definiteness from the second DP, and it results in a nested adjectival modification. The syntactic analyses proposed in Siloni (1996, 1997) for these constructions and adopted in this dissertation derive both constructions from a similar (though not identical) underlying structure, but they differ in their derivation and resulting surface structure. It is widely assumed in the literature that the unique phonological properties of the CS follow from its status as a single phonological word (see Chapter 2, §2.5). The phonological differences between the two complex nominals have consequences for the array of prosodic phrasings that are

[146] Recall the Hebrew does not have the Saxon genitive construction like English and Dutch (Brysbaert & Mitchell, 1996). Whether or not a language has the Saxon has been proposed to matter in processing of Relative Clauses (Frazier & Clifton, 1996); see further discussion in §8.3 below.

hypothesized to be associated with each. If there is a close relation between prosodic phrasing and attachment interpretations, then the prosodic difference could in turn differentially affect RC-attachment for these two complex nominals. The three experiments presented in this dissertation, taken together, reveal aspects of the nature of the interface between prosodic phrasing and syntactic/semantic structure for each of these constructions.

In addition to the variation in the complex nominal to which the RC is attached, the experiments also manipulated the length of the RC. RC length has been shown to affect ambiguous attachment interpretation (see Chapter 4 for references). For silent reading, the general observation is that a short constituent is more likely to attach at a low site, and a long constituent is more likely to attach at a high site (the 'anti-gravity' effect; Fodor, 1998). This variation has a natural explanation in terms of prosodic phrasing, on the assumption that readers project a prosodic contour onto a written sentence. For RC-attachment, the explanation is as follows: A long RC, long enough to constitute a prosodic unit on its own, is likely to trigger a prosodic break before it, thus separating it from the preceding material in the sentence. This discontinuity in the prosodic signal, according to the IPH, would be interpreted by the parser as signaling a discontinuity in the syntactic structure (since the parser applies the alignment principles of the grammar), and thus the RC would be interpreted as attaching high and modifying the whole complex nominal. A prosodic break before a short RC is less likely on phonological grounds, so a short RC is likely to be grouped as part of the preceding material in the sentence, which would amount to being attached low: if N2 is grouped with the RC into one prosodic unit, the low attachment would be the result of prosodic structure; if the whole complex

nominal is part of the same prosodic unit as the RC, the RC may be interpreted as attaching low for locality reasons. This dissertation therefore also examined, for each of the complex nominals under investigation, whether variation in length of the RC results in different prosodic phrasing for this construction, with consequences for attachment interpretations of the RC.

8.2 Summary of experimental findings

The data from the three experiments reported in this dissertation show clearly that the interpretation of FS sentences can be explained in terms of the prosody projected onto the sentence: the overt prosody projected in reading out loud differs for the experimental conditions tested here in revealing ways; the interpretation of overt prosody in listening varies as a function of the prosodic structure of each sentence; the implicit prosody assumed to be projected in silent reading is interpreted in a similar fashion to the way the overt prosody is interpreted in listening. The data for the CS are more complex in that although prosody is shown to play a role in affecting attachment interpretation, it is not the only factor affecting interpretation. Other factors, which are possibly related to the intrinsic properties of the CS, seem to override the effect of prosodic phrasing in interpreting the RC in CS sentences. In what follows, I present a summary of the findings for each of the complex nominals investigated in this dissertation.

8.2.1 The Free State

The FS data from the three experiments reported in this dissertation are highly compatible with each other, as discussed in §7.5 of Chapter 7. The silent reading study

(presented in Chapter 7) compared attachment preferences for short and long RCs and revealed that RC length has an effect on the interpretation of the RC: long RCs were more likely to attach higher than short RCs. This evidence supports the hypothesis that FS is driven by implicit prosody. The production study (reported in Chapter 5) elicited productions of the RC construction for short RCs and for long RCs, when contextually disambiguated for high attachment and for low attachment. The data from the production study indicated that for long RCs and for high-attaching RCs, a pre-RC break is dominant; short RCs attaching low were the only cases in which variation in prosodic phrasing was observed. Based on the production data, it was possible to conclude that the default prosodic pattern for long RC sentences in Hebrew is one in which the RC is preceded by a prosodic break. For short RCs, it was not possible to make a prediction as to the default prosodic phrasing. The listening study (reported in Chapter 6) examined, for each of the prosodic patterns observed in the production study, what the resulting attachment interpretation was for an ambiguous RC sentence and found that attachment varied as a function of prosodic phrasing, with more high attachment interpretation for pre-RC breaks than for the other two patterns observed in the production study, thus confirming the predictions that a prosodic break is interpreted as discontinuity in the syntactic tree, while the lack of prosodic break (and grouping of the RC with preceding materials in the sentence – i.e., N2 or both nominals) gives rise to locality considerations, consequently driving the attachment low.

Taking the results of the three experiments together, it was possible to explain the attachment facts as following from the prosodic facts, by assuming the proportion of times in which a prosodic structure was predicted to occur in silent reading for each of

the RC length conditions, and the likelihood of a given attachment interpretation for that prosodic structure: For long RCs, Phrasing 2 was the prosodic phrasing computed for the majority of sentences. Following the listening study, it was possible to assume that 55.1% of the times,[147] long RC sentences would be interpreted as attaching high, which accounted for the 48% high attachment interpretation found for long RCs in silent reading. For short RCs, no default prosodic pattern was found in the production study. Assuming, following the listening data, that each phrasing pattern is associated with a different attachment interpretation, different combinations of the proportions of each of the phrasing patterns were considered in order to calculate the predicted attachment rate for short RC sentences, and it was possible to arrive at a predicted low attachment interpretation of the short RC quite similar to the low attachment rate observed in the attachment study.

To summarize, the results of the three experiments reported in this dissertation support an IPH-based explanation of the attachment facts for sentences containing FS nominals. While the effect of implicit prosody cannot be tested directly, it is possible to infer the role of implicit prosody in silent reading based on the overt production of contextually disambiguated sentences and the interpretation of the observed prosodic phrasing in listening. Based on the listening and production data, it was possible to arrive at the phrasing pattern(s) assumed to be computed for each RC length, and the predicted resulting attachment interpretation in silent reading.

[147] Note that the observed 55.1% high attachment for Phrasing 2 sentences is for short RC sentence; long RC was not tested for FS sentences in the listening study. However, it is assumed that the attachment rate found for short RCs preceded by a prosodic break would be similar to that for long RCs. See Chapter 7, §7.4 for arguments supporting for this assumption.

However, the Hebrew data is puzzling when compared to data from languages like French and Bulgarian in silent reading studies. Before the experimental project began, it was anticipated that Hebrew would be a high attaching language, similarly to French and Bulgarian. Fodor (2002) proposed that French shows high attachment in silent reading because it has a high rate of breaks before the RC in production. If this is true, then it should apply also in Hebrew. At present, I have no explanation for this, but it will be a focus of future research. For example, f_0 might be a significant factor, but there is no data establishing its relevance to RC attachment in either French or Hebrew.

8.2.2 The Construct State

In contrast to the length effect found in the FS, the silent reading data for the CS did not find a significant difference in the interpretation of short and long RC, although numerically, more high attachments associated with a long RC than with a short RC. This lack of length sensitivity is expected, given the listening data, as discussed below. The production experiment confirmed the assumption widely made in the literature about the CS constituting one phonological word, and showed a marginal use of a phrasing in which a break occurs between the nouns (Phrasing 1). The production experiment also showed, similarly to the FS data, that a pre-RC break was the predominant pattern for long RCs, regardless of attachment interpretation, providing further support for Phrasing 2 as the default phrasing of long RCs in Hebrew, and that high attachment triggered a pre-RC break (almost) regardless of length considerations. Short RC attaching low was the only case which does not fit one of the two sources of pre-RC break. Since the choice

of phrasing pattern for short RC cases varied by attachment height, it was not possible to determine a single consistent phrasing pattern associated with short CS sentences.

The listening data found that both Phrasing 0 (with no prosodic break in the *N1 N2 RC* sequence) and Phrasing 2 (with a pre-RC break) yielded high attachment interpretations. Although the difference between the interpretations of the two phrasing patterns was significant, the effect was reduced compared to the equivalent FS cases. In particular, the high attachment interpretation of CS sentences with Phrasing 0, was contrary to the predicted low attachment rate (following from Recency considerations). This suggested that other factors must be at play in CS sentences, driving attachment interpretation high and overriding prosodic effects. Furthermore, overall attachment in the CS was found to be higher than in the FS, suggesting a high-attachment tendency in the CS that is unrelated to manipulations of length and prosodic structure.

The listening study also found, similar to the findings of the FS data, that pre-RC breaks were interpreted by listeners as high attachment. This was found to be the case for both short and long RC preceded by a prosodic break. Remarkably, the rate of high attachment was nearly identical for both length conditions, suggesting that a prosodic break before the RC is interpreted as indicative of a break in the syntactic structure, regardless of the length of the RC. But since the CS was found to be less sensitive to prosodic and length manipulations, the CS data by itself cannot be taken as absolute support for this claim.

These findings support a view of the CS as a construction predisposed to the high attachment interpretation. This high attachment tendency seems to override RC length manipulations or the implicit prosody presumably computed in silent reading. As

mentioned above, what needs explaining is the surprisingly high attachment rates in the listening data; the silent reading data follow from the listening data and therefore are not surprising in and of themselves. §8.3 of Chapter 8 presents several proposals which are considered as explanations for the high attachment of the RC in the case of the CS.

8.2.3 Major conclusions based on experimental findings

To summarize, the findings in this dissertation lead to two major conclusions: one is that prosody, both explicit and implicit, affects attachment interpretation in listening and in silent reading; this conclusion stems from the findings for the FS construction. Based on the production experiment, which revealed the prosodic phrasing associated with the different sentence types investigated in this study, it was possible to predict the attachment facts in silent reading from the listening data.

The second main conclusion of this dissertation is that some constructions, such as the CS construction, are inherently biased in terms of their attachment preferences, in which case even overt prosodic effects are reduced or eliminated, as are other prosody-relevant manipulations such as the length of the RC. It would be interesting to discover whether there are other such constructions in Hebrew or in other languages.[148] This finding about the relatively inflexible interpretation of the CS was supported by the higher attachment interpretation of the CS compared to the FS, for both prosodic phrasings that were tested in the listening study and also the fact that those two prosodies yielded more similar proportions of high attachment for CS, as compared with the FS. More generally, this implies that the interpretation of some constructions can be

[148] Lehiste (1973) noted that "deep ambiguities" where the two analyses do not differ in surface structure are not susceptible to prosodic disambiguation, but of course this is a very different situation from the Hebrew CS. See Abdelghany (in preparation) for data on prosody-syntax interface for CS Standard Arabic.

independent of the prosody-syntax interface, such that even when prosody is overt in the stimulus, the perceiver's interpretation is determined in large part by non prosodic properties of the construction. Again the question arises whether this is an isolated property of the Hebrew CS or is exhibited in other constructions in other languages. Some suggestions are given in the following section as to what the specific nature of this bias might be, but at present it is mainly unexplained.

8.3 Addressing the CS puzzle

This section explores what might be affecting the overall tendency for high attachment interpretation in the CS data in listening. Before entertaining a few hypotheses which might account for this, let us first consider some explanations which might be invoked, but which turn out not to account well for the data. Hebrew provides us with a contrast between the FS and CS constructions, which is particularly striking because the two constructions are perfectly matched in other respects. This makes it possible to isolate the factors that may be at work in affecting RC interpretation in one construction but not in the other, while avoiding cross-language differences which may independently affect interpretation. Consider first the unique properties of the CS. On the one hand, N1 of the CS is the syntactic head, thus more salient for RC modification. However, despite its head status, N1 of the CS is otherwise weak – it does not carry a phonological main stress and in fact is phonologically in a reduced (construct) morphological form; in its construct form, it is also syntactically restricted in several ways (e.g., no adjectival modification or definiteness marking), as discussed in Chapter 2, §2.2. N2, on the other hand, heads a full external projection syntactically and

phonologically it carries a main stress; in fact, N2 carries the main stress of the whole CS, thus is viewed as the phonological head of the complex nominal (Bat-El, 1993). The unique properties of the CS thus seem to present a clash between two pressures: the syntactic head and the phonological head, each pulling RC attachment to it.

Consider an explanation of the CS facts based on the 'Focus Attraction Hypothesis' (Shafer et al., 1996); this is another prosodic account, but it was paid little attention so far. This hypothesis proposes that "In sentences where either of two NPs may head the relative clause, placement of a pitch accent on a noun can focus the NP headed by that noun, increasing the likelihood of the perceiver's attaching the relative clause to that NP." (Schafer et al., 1996:136). Thus, the RC is attracted to the focal-stressed head noun. In their study of RC sentences in English, Schafer et al. (1996) found that the word carrying the most prominent pitch accent is interpreted as conveying the focus of the sentence and thus attracts RC modification. Although a stressed head can be seen as focused, it may not have the same effect as a focal element, since the focus accent studied in Schafer et al. (1996) is different from the rule-driven lexical stress in the CS. Nonetheless, it is interesting to try and extrapolate from the Focus Attraction Hypothesis the predictions for the Hebrew CS. In the case of the CS, since N2 is the phonological head, and in fact the only accented word within the complex nominal, the hypothesis predicts that it would be interpreted as the focused word and attract the RC, resulting in a low attachment interpretation of the RC. However, clearly this is not the case in the Hebrew CS: the findings of both the listening and silent reading experiments suggest that N2 does not attract RC modification; on the contrary, the CS exhibits an overall bias

towards N1 (high) attachment interpretation of the RC, even though N1 does not carry main stress and therefore cannot be focused.

Hemforth and colleagues (Hemforth et al., 2002; Hemforth et al., 2006) propose an explanation of attachment preferences in ambiguous construction based on an informational approach. In particular, two aspects are discussed: (1) richness of the attachee and (2) prominence of the host. Each is discussed in turn. The 'richness of the attachee' proposal (Hemforth et al., 2006; see also discussion in Chapter 6, §6.3) offers a pragmatic account of the length effect, which assumes that a long constituent, which contains more information, requires more processing effort; this enhanced processing effort is justified if the long constituent modifies a more central element in the discourse. Thus a higher attachment tendency, to more prominent elements, is predicted for more informative constituents. Note, however, that in constructing the experimental sentences used for this dissertation, careful attention was paid to ensure that the RC length does not add crucial information (e.g., *še-hištaala* 'who coughed' vs. *še-hištaala ši'ul kaved* 'who coughed a heavy cough'; see also Appendix A-1 for a complete list of materials); therefore this is not a plausible explanation for the results of the present experiment. Also, this proposal does not explain the absence of a length effect in the CS and its presence in the FS data.

Hemforth et al. (2002) propose that the prominence of the host affects attachment interpretation: "focused phrases attract modification while non focused phrases resist it" (p. 22). They investigate position effects in the RC construction and show for German and Spanish RC sentences (but not for English) a position effect: When the complex NP + RC was in preverbal subject position (assumed to be associated with topics, which are

generally treated as 'old' information and not focused), high attachment was reduced compared to when it was in post verbal position (assumed to be a position for asserted or focused information).[149] In addition, Hemforth et al. note that focus on an NP can be interpreted as focus on the head of the NP, N1; as has been shown by Schafer et al. (1996), focus attracts RC modification and thus "the greater tendency of RCs to modify N1 in object position than in subject position may be attributed to an implicit focus effect" (Hemforth et al., 2002:21-22). While Hemforth et al. (2002) explain the cross-linguistic difference based on informational structure, Fernández and colleagues (Bradley, Fernández & Taylor, 2003; Fernández, Bradley, Igoa & Teira, 2003; Fernández & Bradley, 2004) found prosodic correlates for the position-sensitive attachment data reported in Hemforth et al. (2002), which could explain the attachment data in terms of processing at the prosody-syntax interface. As an explanation for the high attachment bias in the CS data, however, the informational prominence explanation is unsatisfactory, along similar lines to the prosodic prominence explanation of Schafer et al. (1996), in that in order to explain the CS data, it would predict that N1 would have prominence. But in fact prominence cannot be placed on N1 in the CS (it is not possible to translate information prominence as prosodic prominence, since N1 of the CS is a bound form which does not carry main stress), and therefore it cannot explain the high attachment in the CS.

Another potential explanation links the CS to the closely related structure of Hebrew compounds. Compounds in Hebrew share some important properties with the

[149] In a summary of her experimental findings from different works, Hemforth (2006) reports a study by Shuval & Hemforth (2006), which found a similar position-sensitive pattern for Hebrew translations of the materials used for their German, English and Spanish studies (presumably Hebrew FS was used for the comparison). I was not able to obtain further information about the abovementioned study by Shuval and Hemforth.

productive CS, as discussed in Chapter 2, §2.4. A crucial difference between compounds and CS nominals, regardless of their detailed analysis, is that N2 of the compounds cannot be modified, while N2 of the CS can be modified by an adjective or an RC. A possible explanation for the high attachment tendency of the CS may be that the CS is interpreted as a compound, thus blocking the N2 interpretation and creating a bias for N1 interpretation. However, the silent reading data found that the average tendency for high attachment in the CS is 60%, with 40% attachment to N2. If CS sentences were interpreted as compounds, a more categorical rejection of N2 modification would be expected, leading to an extremely high attachment of the RC, but this was not the case. The 40% of CS cases which had a low attachment interpretation confirm that N2 modification was possible at least for many of the target items.[150] It should be noted that the materials for the experiments presented in this dissertation were tested to ensure that both N1 and N2 modification was possible, and that they were equally acceptable in the RC attachment assessment, reported in Chapter 4, §4.3. This assessment test confirmed that there is no semantic bias towards N1 modification in the materials used for these experiments. The RC attachment test further confirmed that N2 modification is as acceptable as N1 modification for both the CS and the FS. The analysis compared FS and CS cases with identical content, and found no difference in the acceptability of N1 and N2 modification between the two complex nominals, thus also rejecting potential explanations of the high attachment of RC in the CS cases based on semantic, pragmatic or syntactic inherent bias towards N1 interpretation in the CS only. The complex nominal rephraseability assessment reported in Chapter 4, §4.2 ensured that all the materials can

[150] There was no individual experimental item in which low-attachment was exclusive to CS (unlike FS); if a specific item was interpreted preferentially with low attachment, it was so for both CS and FS instantiations.

be used in both the CS and FS forms, and this confirms that indeed the CS that were used are productive, since compounds cannot be paraphrased into FS forms (see Chapter 2, §2.4). Another test for determining whether a complex noun is a compound or a CS is offered by Rosén (1977): CS, but not compounds, exhibit the replaceability of any one of the components of the complex nominal by members of its substitution class (as noted in Chapter 2, §2.4 of Chapter 2). For example, for the CS *me'amen ha-mit'agref* 'the coach of the wrestler', N1 (*me'amen* 'coach') can be substituted by *ozer* 'helper', deriving *ozer ha-mit'agref* 'the helper of the wrestler' and N2 *(mit'agref* 'wrestler') can be substituted by *saxyan* 'swimmer' to derive *me'aen ha-saxyan* 'the coach of the swimmer'. This test further confirms that the experimental items used in this study are not compounds. Finally, based on intuition, several Israeli linguists I consulted strongly rejected the proposal that CS sentences are read online as compounds; Hebrew speakers intuitively know which complex nominal is considered a compound and which is a productive CS, presumably as part of their knowledge of the Hebrew lexicon (Berman, 1978; see further discussion in Chapter 2, §2.4).

An interesting contrast arises if one considers solely the number of prosodic words in experimental sentences containing the CS as compared to those containing the FS. Since the CS is considered a single prosodic word (as opposed to the FS which contains 2 prosodic words), sentences containing the CS have 1 prosodic word less than those containing the FS. The fact that CS is lighter makes it less likely that it would induce a prosodic break following it and preceding the RC. This would predict less high attachment in the CS than in the FS, based on the prosodic differences between the two complex nominals. However, as the production experiment shows, pre-RC breaks in the

long RC conditions are equally likely for the CS and FS and more likely in the short RC conditions for the CS than for the FS (presumably due to the marginal use of the *N1)(N2 RC* phrasing in the CS). In addition, both the listening and silent reading experiments show more likelihood of high attachment in the CS than in the FS. Therefore, this explanation does not seem to explain the high attachment tendency in the CS, though it does provide a reason for why length effects might be suppressed in the CS.

Another type of explanation which can be explored is the 'Gricean' account invoked in Frazier & Clifton's (1996) Construal hypothesis. According to this hypothesis, one of the interpretive conditions that influences perceivers' choice between several potential hosts for the relative clause is Grice's conversational maxims, in particular the maxim "avoid ambiguity" under the category of Manner (Grice, 1975: 45-46). This maxim requires speakers to be as unambiguous as possible and therefore in cases where there is an unambiguous structure available, they expect speakers to use it rather than using an ambiguous structure. In English, an RC following the Saxon genitive ('N2's N1') unambiguously modifies N1 and therefore perceivers will assume that it will be used where N1 is the host, while the alternative Norman genitive construction ('N1 of N2') will be reserved for when N2 is the host. This hypothesis would thus predict that for languages like English, which have one unambiguous and one ambiguous construction, the interpretation of the RC in the ambiguous form would be biased by the existence of the alternative unambiguous form. Note that this hypothesis was disproved in Lovrić (2003), but Hebrew might present a new case for examination. It is conceivable that the FS and CS would inevitably tend to associate one attachment to one construction and the other to the other construction (even if the tendency is slight). Of the two, the FS is more

likely to associate with low attachment, because the CS cannot have an internal break, which is associated with a very low attachment interpretation, and thus is more likely to be associated with high attachment interpretation. Thus a Hebrew perceiver might expect high-attaching meanings to be conveyed using the CS construction, and low-attaching meanings to be used with the alternative FS nominal. This would explain the overall lower attachment in the FS and the higher attachment in the CS. This is an intriguing possibility; however, it should be noted that the materials for the experiments reported in this dissertation were carefully matched, such that there was no bias in attachment of the RC to one noun over the other of the complex nominal. As noted in Chapter 4.3, an RC attachment assessment confirmed that both N1 and N2 are highly plausible as hosts for the RC, and roughly equally so, both in CS and FS sentences. Additionally, the Gricean account does not explain the insensitivity of the CS to length manipulations, which is an unusual finding. Length effects, which were typically found in prior investigations of the RC constructions in other languages, were only found in the FS; the Gricean account does not explain the contrast between FS and CS with respect to the length factor.

Finally, explanations based on frequency (lexical or construction type) should be considered. Pynte & Colonna (2000) found that the RC tends to attach to a low-frequency noun rather than to a high frequency noun. Thus it is possible that the high attachment bias in the CS stems from a bias in the experimental materials such that the N1 words are less frequent than the N2 words, attracting modification by the RC. While an effort was made to choose words of equal frequency in constructing the materials for the experiments, no word-frequency list for Hebrew was available at the time when the experiment was designed. But regardless of that, the design of the experiments compared

identical material sets for FS and CS sentences, as noted above, and therefore ensured that any residual frequency biases would appear in both the FS and CS cases. Since the high-attachment bias was found only for the CS cases, it is safe to conclude that it does not stem from lexical frequency bias towards N1 in the materials for this experiment.

Another type of explanation focuses on the low frequency of the CS construction as a whole. Ferreira et al. (1996) found that low frequency words presented auditorily in a sentential context took longer to process than high frequency words. This is tantamount to inserting a prosodic break after the low-frequency word, which can be seen as facilitating processing and reducing processing load. The CS as a whole can be viewed as a low frequency entity: it is generally assumed that the CS is mainly used in high registers and in written language, in contrast with FS cases which are used more often in informal context, as discussed in Chapter 2, §2.1. Thus, it is possible that a prosodic break is placed after the low-frequency CS, resulting in a high attachment interpretation of the RC in silent reading. The results of the production data for the CS indeed show that three of the four CS conditions (high- and low-attaching long RCs and high attaching short RCs) are uttered with a prosodic break after N2. However, low-attaching short RCs were produced both with a break after N2 and with no break in the complex NP + RC sequence. The difference between short and long RCs found in the production study did not manifest itself in a difference in attachment preference for short and long RCs in the listening data; not only Phrasing 2 was interpreted as attaching high, as expected, but also Phrasing 0, a puzzling finding, as discussed above. Therefore, it seems that the high attachment bias in the CS is not due solely to the presence of a prosodic break after N2, but to other factors as well.

Keeping with the frequency-based explanations, a probabilistic model which can be explored with relation to the CS data is the Tuning Hypothesis (Brysbaert & Mitchell, 1996; Mitchell, Cuetos, Corley & Brysbaert, 1995). The Tuning Hypothesis proposes that the relative frequencies with which high and low attachments have been encountered in comparable structures in the past exclusively affect the initial parsing choices for the current ambiguous input. If corpus data show there are an equal number of occurrences of short and long RCs attaching high in the CS, then the high attachment in the CS can be predicted by the frequency-based model. However, the model does not address the question of why the CS is interpreted more often with high attachment than with low attachment. This model would also predict that in the FS, long RCs would have more occurrences of high attachment than short RCs (see also footnote 143 in Chapter 7); however, the model does not provide an explanation based on the different properties of the FS and CS to account for the reason why the FS behaves differently compared with the CS.

The explanations above fail to capture the peculiarity of the CS data, either because they predict a low attachment preference for the CS (clearly counter-evidenced by the data of the experiments) or because they make predictions which should apply to both CS and FS cases (which was not the case). Any hypothesis that might clarify the source of the high attachment bias in the CS data needs to focus on an aspect of the CS that is unique to this construction, and can explain the different behavior it manifests in contrast to the FS. Let us explore two remaining hypotheses that may shed light on the curious behavior of the CS. One hypothesis involves the information structure of the CS and the other concentrates on its syntactic structure.

The high attachment bias in the CS is sometimes termed N1 attachment, though structurally, N1 modification is actually modification of the whole complex nominal, headed by N1. If CS as a whole is informationally salient (i.e., if the complex nominal as a whole dominates informationally more in the CS than in the FS), then N2, though syntactically modifiable by the RC, cannot be singled out informationally. This would explain the high attachment preference of in the CS.

Retaining as much as possible of the interesting aspects of the approaches on prominence of Schafer and colleagues and Hemforth and her colleagues, we propose here a synthesis: we need to dissociate informational prominence from prosodic prominence in some way. Consider prosodic focus projection: According to Chomsky (1972), the heaviest focal stress in English falls on the last word in the sentence, but can project to constituents larger than the word which is assigned focus (see also Selkirk's (1995) theory of focus; for experimental studies of focus projection in English see Birch & Clifton, 1995 and Schafer et al., 1996). The idea of focus projection is that informational focus projects from the prosodic focus. In the FS, prosodic focus is possible either on N1 or N2, or on the entire complex nominal: if prosodic stress is placed on N1, it renders N1 focused; if stress is on N2, it can be projected as stress either on N2 alone or on the entire complex nominal (i.e., the prosodic stress on N2 is ambiguous with respect to focus: either N2 or the whole complex nominal can be focused). In the CS, on the other hand, it is not possible to focus N1 by placing a prosodic focus on it, since N1 is in its reduced form and does not carry main stress (as discussed above in relation to Schafer et al.'s (1996) and Hemforth et al.'s (2002) proposals). Focal stress can only appear on N2, and it is ambiguous between focus on N2 by itself or focus projection to the entire CS. We

suggest that in the CS, N2 stress is preferentially interpreted as stress on the whole complex nominal (focus of the whole complex nominal is achieved by stressing N2, the location of main stress). If N2 of the CS can be shown to be informationally non-salient, then the tendency for high attachment in the CS (to the whole complex nominal headed by N1) can be explained as following from the salience of the complex nominal as a whole and the lack of prominence of N2 of the CS (as opposed to the salience of N2 of the FS).

The idea that there is an intuitive difference in focus between the CS and FS has been hinted at in the work of Ravid and Shlesinger (1995), as mentioned in Chapter 2, §2.1. They propose that the FS and CS are typically used each to convey a different semantic relation, based on the structural differences between them, which lead to differences in the realization of semantic focus in each construction. The CS is claimed to be treated as a single entity in terms of syntactic operations that apply to it and therefore it is typically used to convey semantic relations with "focus on the head which constitutes the semantic core…, while the modifying component delimits it in various ways." (p.160). Thus it is claimed that the CS is used primarily to convey semantic relations such as classifying relations (e.g., *atar nofeš* 'resort site'), in which the syntactic head (N1, *atar* 'site') carries the main semantic weight and the adjunct (N2, *nofeš* 'resort') narrows its meaning to a 'type of site'. In the FS, on the other hand, there is "a more balanced share of semantic weight between the possessed head and the possessing adjunct", in which two entities are involved, and therefore the FS "expresses possession [e.g. *ha-bait šel ha-melex* 'the house of the king'] more transparently by specifying each

phonologically independent element separately, and by giving explicit expression to the possessive particle *šel* itself." (p. 162).

It is possible to test the status of N2 of the CS in terms of informational salience by posing questions that probe the information carried by the N2 and examining whether the sentence is a legitimate answer to the questions. This would reveal whether N2 alone could be informationally focused in the CS. To check whether N2 in the CS cannot be focused, it is possible to investigate whether a CS can be used in reply to the question in (3). Such a question can in principle be answered either with 'he coaches the wrestler' (4) or 'he is the coach of the wrestler'; the latter can involve either a CS or an FS form, shown in (5). If N2 of the CS cannot be informationally prominent, the CS would not be an appropriate answer to such question, see (5a). In contrast, in the FS in (5b), since N2 can be independently focused, the answer should be appropriate. This is indeed the case according to the intuitions of a number of native speakers consulted. I use the # symbol to indicate an inappropriate answer to the question.

(3) 'et mi hu me'amen?
ACC who he coaches
'Who does he coach?'

(4) hu me'amen 'et ha-mit'agref
he coaches ACC the-wrestler
'He coaches the wrestler.'

(5) a. #hu me'amen ha-mit'agref
he.is coach the-wrestler
'He is the coach of the wrestler.'

b. hu ha-me'amen šel ha-mit'agref
he.is the-coach of the-wrestler
'He is the coach of the wrestler.'

An additional test for the possibility of a narrow focus on N2 in the CS comes from Ladd (1996). Ladd describes an exchange between a reporter and John Dillinger, a notorious bank robber in the 1930's, shortly after his arrest. In answer to the reporter's question 'Why do you rob banks?', Dillinger reportedly replied 'Because that's where the money is.' (Ladd, 1996: 200). While the question was intended with a broad focus, the sarcastic reply treats the question as if it focuses only on 'banks', i.e. interpreting the accent on 'bank' as if it signals narrow focus. Ladd uses this example to provide evidence that the question with a broad focus prominence pattern is potentially ambiguous between a broad and a narrow focus pattern. It is possible to use a version of this example (e.g., 'Why are you a robber of banks?') to test the information prominence in the CS, in contrast to that in the FS. When the question is in the form of a CS (6a), the answer in (7) seems infelicitous (as noted by the # symbol), suggesting that N2 of the CS cannot be narrowly focused. In contrast, when the question is in the FS form (6b), the sarcastic answer in (7) works, suggesting that the focus on N2 in the FS can be interpreted as broad focus on the entire complex nominal or on N2 by itself.

(6) a. #lama nihyeta šoded bankim?
Why became robber banks?
'Why have you become a robber of banks?'

b. lama nihyeta šoded šel bankim?
Why became robber of banks?
'Why have you become a robber of banks?'

(7) ki šam ha-kesef nimca.
because there the-money is
'Because that's where the money is.'

Both the answers in (5) and the questions in (6) show that focusing on N2 in the FS is fine, but it is not possible in the CS. This suggests that in the CS, N2 alone cannot be focused, and therefore it cannot be informationally prominent. Since semantically N2 of the CS is not salient, it cannot bear focus by itself, forcing semantic focus to be inherited upwards. Although prosodic focus can be placed on N2 in the CS, it is interpreted as informational prominence on the entire CS (through focus projection), thus explaining the tendency for high attachment in the CS. This suggests that the FS may or may not be focused, whereas the CS, if it is focused at all, must be focused as a whole (not just N1, and not just N2). If this is the case then the high attachment in the CS may be explained as an inherent requirement in the CS to be a focused constituent, which in turn attracts modification by the RC, regardless of RC length and in spite of prosodic manipulations. The FS does not have restrictions on its information structure (it may or may not be focused), and as has been argued for in this dissertation, prosodic phrasing plays a central role in its interpretation.

It is unclear why the CS should be treated as an informational unit, even though it is semantically compositional (see Chapter 2, §2.1); the following are several speculations. Recall one of the unique properties of the CS, i.e., definiteness spread: the definiteness value of the CS as a whole is determined by the definiteness value of N2 (see Chapter 2, §2.2). Possibly, focus spread is analogous to definiteness spread (a possible link between definiteness spread and focus in Greek is suggested in Kariaeva, 2004). In particular, the notion of shared features may be general in the CS: N2 bears the features on behalf of N1 in the CS (i.e., definiteness and prosodic focus), but although both definiteness and prosodic focus are marked on N2, they spread to the whole

complex nominal. Alternatively, it might consist of standard focus projections (see above), but with a special twist in CS on the assumption that N2 cannot bear semantic focus alone and therefore despite the focal stress being placed on N2, it must be interpreted as focus of the whole complex nominal.

To sum up, the preference for high attachment interpretation in the CS found in the listening data and in the silent reading data suggests that in processing the CS, the information structure limits the possible effect of prosody. The whole complex nominal, headed by N1, is the informationally prominent element in the complex nominal, and although N2 carries the main stress of the CS, N2 cannot be focused and therefore is not likely to be modified by the RC, thus providing a possible explanation as to why attachment to the CS nominal as a whole is preferred over N2 attachment. Of course this proposal is highly speculative at this stage, and further investigation, both linguistically and psycholinguistically is needed.

It is important to consider non-prosodic explanations, such as syntactic ones, as an explanation for the bias in the CS towards attachment to the whole complex nominal. I know of none in the syntactic literature, but a possible idea in order to illustrate this is an explanation that focuses on syntactic differences in RC association between FS and CS.[151] This hypothesis offers a new syntactic analysis of the FS and CS: For the CS, both NP_1 and NP_2 are on right branches, and therefore each is straightforwardly available as a host for the restrictive RC through right association, see Figure 8-1.[152] The relation in the CS structure is one of domination, and since neither NP_1 nor NP_2 is left branching, both are accessible for RC association (note that this syntactic account is different than that in

[151] The syntactic account outlined here was proposed by Marcel den Dikken in informal discussions; errors in representation are mine.
[152] Restrictive RCs association is to NPs only (cf. non-restrictive RCs, which attach to DPs).

Siloni (1996,1997), which is outlined in Chapter 2, §2.5). Importantly, when the RC associates with the projection of N_1 in the CS, it does not strictly speaking associate *just* with N_1: N_1 selects DP_2 and as consequence the projection of N_1 dominates NP_2; therefore, attachment of the RC to NP_1 is attachment of RC to the *entire* complex noun phrase (including NP_2).

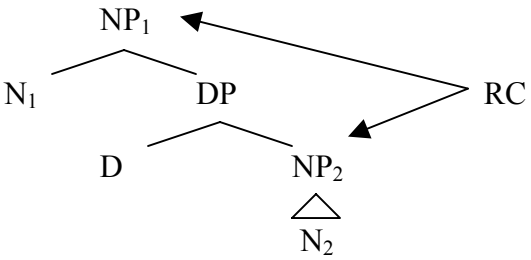


Figure 8-1: Restrictive RC association in the CS

In the FS, on the other hand, while NP_2 is on a right branch, the projection of N_1 is on a left branch: above the NP_1 node there is a functional head projection FP, which is assumed to be the functional projection of the copular element *šel*, following den Dikken's (2005) analysis of the English *of* and the comparable Hebrew *šel* as copular elements arising from a lexicalization of a linker head via copular inversion of the subject (N_1) and the predicate (N_2). Whether or not this view is adopted, it is certainly not the case that *šel* is an NP_1 projection. The structure for the FS is presented in Figure 8-2.

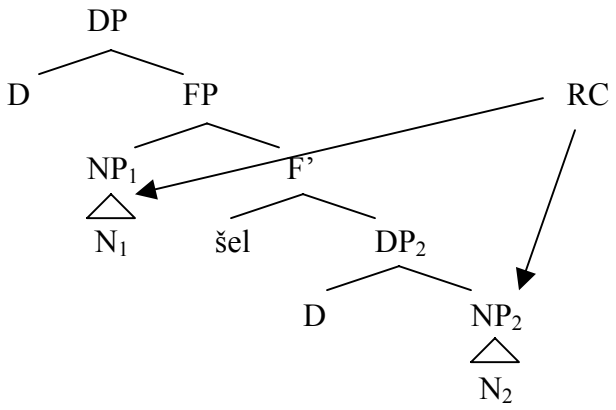


Figure 8-2: Restrictive RC association in the FS

Since the projection of N_1 is on a left branch, it is not directly associable to a relative clause that linearly follows N_2. High attachment of the RC could only take place by extraposing the RC (the Linear Correspondence Axiom of Kayne (1994) prohibits rightward movement but the effect of RC-extraposition can be achieved by means of leftward movement). Although both attachment sites are possible syntactically, low attachment of the RC (association to a right-branching node) is easier to construct, compared to high attachment (association to a left-branching node), and therefore is easier to process; structurally, left association of the RC to NP_1 is more difficult than right association of RC to NP_2 is. According to this analysis, then, RC association in the CS should be as easy for both attachment sites, NP_1 and NP_2, since both are right branching. In the FS, on the other hand, RC modification of NP_2 (which is right association) should be easier than that of NP_1 (left association of the RC), unless disambiguated by context, semantics, or syntactic factors such as number and gender.

If, in the CS, the RC can associate either with the projection of N2 or with the entire CS (including N1), but not with the projection of N1 by itself (to the exclusion of the projection of N2), then whenever the RC is interpretively associated with N1, it

syntactically associates with the entire CS. In the FS, on the other hand, association to N1 alone should be possible. One way to test the hypothesized structures in Figures 8-1 and 8-2 is to look at Principle C reconstruction effects. According to Principle C of the Binding theory, R-expressions must be free; therefore, they cannot be bound (c-commanded) by an antecedent. In the surface structure of the sentences in (8) there is a gap inside the RC of the head that moved (*mit'agref*, 'wrestler'), and the moved head is not c-commanded by the pronoun *hu* ('he'). When the sentence is interpreted at LF, the underlying structure is reconstructed (the head is placed back to the site from which it was moved) in order for the binding relation to hold. If in the underlying structure, the pronoun c-commands the head *mit'agref* ('wrestler'), Principle C reconstruction effects will emerge.

(8) a. ra'iti 'et ha-'axot šel ha-mit'agref$_i$ še-hu$_i$ haxi 'ohev
saw(1st sg) ACC the-sister of the-wrestler that-he most likes

?? b. ra'iti 'et 'axot ha-mit'agref$_i$ še-hu$_i$ haxi 'ohev
saw (1st sg) ACC sister the-wrestler that-he most likes

Both: 'I saw the sister of the wrestler that he likes the most', with the intended meaning: 'I saw the sister of his that the wrestler likes the most'

In the FS sentence in (8a), *mit'agref* can be co-indexed with *hu*. Therefore it seems that there is no effect of Principle C reconstruction in the FS. We do seem to have such effect in the CS sentence in (8b), or at the least the sentence is less acceptable with coindexation of *mit'agref* and *hu*. This suggests that in the CS, in reconstructing the head of the RC into the gap inside the RC, the whole complex nominal (including N2) is reconstructed and therefore the pronoun c-commands N2, giving rise to a Principle C effect of the binding theory. This supports the idea that in the CS, *interpretive* association

of the RC with N1 really involves *syntactic* association of the RC with the entire CS (including the projection of N2). The FS, in contrast, shows no Principle C reconstruction effect. The absence of Principle C reconstruction effects in (8a) suggests that in the FS, it is possible for only N1 to be reconstructed into the gap; this is compatible with the structure suggested: If the FS has a larger structure to which RC attaches with a right branching association, then we would have a Principle C reconstruction effect in the FS, which we do not find.

If this was the complete story and no other factors were involved, then we would expect low attachment in the FS for purely syntactic reasons. But of course other factors are involved in processing. This explanation does make low attachment more available in the case of the FS, since its structure is easier to construct; this view is compatible with the experimental findings of this dissertation. While the analysis does not directly explain the CS data, it does suggest that there should be no preference for N2 attachment in CS, as opposed to the N2 preference in FS. While this does not explain the high attachment found in the CS data, it is compatible with the explanation of the CS as informationally salient as a whole, as discussed above. Its should be noted, though, that if the FS in Hebrew is syntactically like the 'N1 of N2' in French, German, Croatian etc. (which might not be the case), and if no other factors other than syntax are at work, then it would predict a similar lack of high attachment and no length effect in these other languages, contrary to facts.

8.4 Future directions

The following are clearly issues that would benefit from further investigation. One avenue of research is the ranking constraints on prosodic phrasing in Hebrew. This dissertation provided new data on prosodic phrasing in Hebrew at the sentential level, a field which has largely been unstudied in Hebrew. Shaked (2007) used the FS data from the production study reported in this dissertation to determine the ranking of constraints on prosodic phrasing in Hebrew and showed that in Hebrew, the alignment constraint is ranked higher than the length constraints. Furthermore, the data showed that the two length constraints, *BinMin* and *BinMax*, are unranked with respect to each other, indicating that Hebrew in general prefers phrases optimally of 2 prosodic words in length.

While the ranking of the alignment and length constraints has been strongly supported by the Hebrew data on the phrasing of the complex nominal + RC region, it would be valuable to further evaluate the proposed ranking with elicited data on other ambiguous constructions for which different constraints are relevant, or in other positions in the sentence. Additionally, data need to be collected on the effects of the alignment and length constraints (among other constraints on prosodic phrasing) in fully spontaneous speech. The proposed ranking was determined based on the semi-spontaneous elicitations of the production study. In my judgment, this is the most natural way to elicit utterances in a controlled experimental setting. The ambiguous RC construction with two nouns is unfortunately not common in spontaneous speech and therefore it is difficult to collect such data from spontaneous conversations. However, it is possible to collect speech from spontaneous conversations and assess the ranking of the

constraints based on those productions, to check whether they support the ranking needed for explanation of the RC-attachment facts.

Other questions invite further study with respect to the ranking of constraints on Hebrew prosodic phrasing. One interesting finding that needs further exploration is the frequent occurrence of a break after the verb in the elicited utterances of the production study. The break after the verb is not predicted by the alignment constraint, and yet it occurs in the majority of the cases, and intuitively sounds natural for the sentences produced in this study.[153] Assuming the high ranking of the alignment constraint in Hebrew, it is likely that a break would occur after the subject in order to satisfy the alignment constraint. However, the Hebrew data revealed a case of competition between a break after the verb and a break after the subject, with very few cases in which a break occurred both after the subject and after the verb. This suggests that some other constraint is ranked higher than the alignment constraint in Hebrew (e.g., eurhythmic constraints such as Uniformity: " a string is ideally parsed into same length units" (Sandalo & Truckenbrodt, 2002: 298)); the interaction of these constraints may be what is shaping the phrasing structure found for the RC sentences in Hebrew (i.e., the break after the verb). One way in which the investigation may be revealing with respect to the post-verbal break may be an examination of whether the tendency to break after the verb is reduced when the RC sentence is used with a covert subject ('little pro'). In such sentences the verb is the first element in the sentence, therefore a break after it is less likely; this may provide clues as to the type of constraint that outranks alignment considerations in the

[153] A break after the verb is uncommon in general, but it was also reported as a common phrasing pattern in SVO sentences with long objects in Catalan (Prieto, 2005; D'Imperio et al., 2005) and in interrogative questions in Georgian, where the word order is S-V-Complement (Jun et al., 2007). Bachenko & Fitzpatrick (1990) report cases in which a break occurs after the verb in English, among other phrasing patterns they found, which vary according to length and word frequency.

materials studied in this dissertation in favor of a post-verbal break. Varying the overall prosodic weight of the sentence, by using covert subjects or heavier subjects (those consisting of nouns and modifiers) might also prove revealing in the exploration of the post-verbal break.

Additionally, the elicited production data reported in Chapter 5 of this dissertation was generated using a ToBI-like coding system, designed specifically for the purpose of analysis of the data for this dissertation. This ToBI-like system, as described in Chapter 5 §2.5.1, is rudimentary at this stage, and is based on the scarce literature on prosody at the phrasal level currently available for Hebrew (see Chapter 3). However, the instrumental analysis of duration and f_0 data (in §2.5.2 of Chapter 5) support the findings of the ear-judgment phrasing data, and provide confidence that the coding system used is feasible and taps into prosodic breaks of the relevant sort (assumed to be intermediate phrasal breaks). For the future, more research will be needed to develop a thorough characterization of the prosodic levels in Hebrew and of the different break index types, based on systematic instrumental evidence.

The most interesting points to pursue are related to the discovery about RC attachment to the CS, which opens many new questions for future research. In particular, it will be important to uncover what inherent property of the CS drives the interpretation high, overriding length manipulations in the attachment data and reducing the effect of prosodic phrasing, as was observed in the listening data. The proposal that the CS as a whole is informationally salient, while N2 by itself is not prominent and therefore is more likely to reject modification by the RC, as suggested in §8.3, deserves further investigation. This can be done by further modification of N2, which may make it a more

prominent host for attachment of the RC (Thornton, MacDonald & Arnold, 2000). In Hebrew, adjectives generally follow the nouns they modify. For CS nouns, adjectival modification is not useful for this purpose because an adjective modifying N1 appears linearly after both N1 and N2 of the CS (and under certain circumstances may also ambiguously modify either of the nouns), as discussed in Chapter 2. Therefore, an adjective modifying either of the nouns of the CS would follow the complex nominal and consequently may be ambiguous between N1 and N2 modification, just like an RC. Numerical modification, on the other hand, precedes the noun, and therefore unambiguously modifies only one of the nouns in the CS. The example in (9) illustrates this.

(9) dodey šney ha-yeladim še-‘azvu
uncles two the-children who-left$_{pl}$
‘The uncles of the two children who left.’

The potential effect of numerical modification of N2 on RC attachment preference deserves experimental investigation; meanwhile, native speaker intuitions have suggested that it would induce a break internal to the CS, thus weakening the phonological tightness between the two nouns in the CS. A prosodic break between N1 and N2 may reduce the likelihood of a break immediately following N2, and therefore make high attachment interpretation in such cases less likely. So this may be a case where the requirement for phrasing of one prosodic word in a single phrase can be overridden by other factors. RC attachment to CS sentences in which N2 is numerically modified may contribute to an understanding of the forces driving attachment high in the sentences studied in this dissertation.

Also of interest for further investigation is the phonological status of the CS. As largely assumed in the Hebrew literature, the CS is a single phonological word; this was supported by the production data reported in Chapter 5, which show that a break within the CS is highly unlikely. Intuition suggests that several manipulations could make a prosodic break within the CS more likely. One is to create a CS with a string of nouns. As mentioned in Chapter 2, one of the properties of the CS is recursiveness: several nouns can compose one CS, as illustrated in example (10).

(10) mazkirat yošev roš irgun hamorim
secretary sitting head union teachers
'The secretary of the chairman of the union of the teachers'

Long sequences of nouns in the construct are uncommon, but a sequence of 3 nouns seems reasonably productive, according to intuition. It is interesting to see whether in a sequence of 3 nouns, a prosodic break is more permissible within the CS, despite the fact that theoretically, primary stress appears on the last noun in the sequence only, suggesting its status as a single (though quite long) phonological word. If indeed a break between two of the three nouns in a recursive CS is possible, then this could be manipulated experimentally in order to examine whether a CS internal break reduces the likelihood of high attachment interpretation of an RC. Intuition suggests that a break within the CS would lower attachment in the CS similarly to its effect in the FS cases studied in this dissertation.

As discussed in §8.3, one of the interesting topics for further investigation is the placement of focus in the CS and the ways in which focus is realized on the entire CS, if possible (focal stress was not an investigated factor in this dissertation, though it has been

shown elsewhere to affect attachment preferences, see Schafer et al., 2000b). If N1 does not carry a main stress, how is focus on N1 realized? The possibility was raised that focus is marked on N2 but N2 by itself cannot be focused; rather, focus, quite like definiteness marking in the CS, spreads to the whole complex nominal. Further study of the phonological properties of the CS will be needed in order to determine how focus is realized in the CS, and its effect on the processing of this complex nominal. Another interesting idea which is compatible with the focus projection hypothesis and needs more attention is the syntactic structure proposed for the CS and FS in §8.3. Together, these two hypotheses may advance our understanding of the bias in the CS towards attachment to the whole complex nominal headed by N1.

The listening data raise two additional issues. The first relates to the naturalness of the stimuli. For each complex nominal, the phrasing patterns associated with it in the production study were tested in the listening study. For long RCs, it was shown that one phrasing pattern is overwhelmingly used for all conditions, suggesting that it is the most natural way of producing long RC sentences. For short RCs, on the other hand, it was not possible to determine according to the production data which is the most natural prosody for short RC cases. The listening study used a variety of phrasing patterns, but it is possible that among the phrasing patterns used, one is the most natural prosody for this construction. It is inevitably difficult to establish the default prosody for an ambiguous construction except in the special case where both of its readings favor the same prosodic pattern (as was shown to be the case for the FS long RC). So I must leave it open what default, if any, holds for the short RC.

One final possible avenue for research of the RC construction in Hebrew is a cross-linguistic study of Hebrew and Standard (Egyptian) Arabic, two Semitic languages which despite close resemblance in several language areas, seem to differ in their prosodic characteristics. Abdelghany (in preparation) investigated the RC construction in Standard Arabic and found that Arabic tends to use long prosodic phrases (as was also found for Cairene Arabic in Hellmuth, 2004) as opposed to the findings in this dissertation, that Hebrew prefers short phrases (see §5.4 in Chapter 5, which further comments on the differences in the hypothesized constraint-rankings for Hebrew and Standard Arabic). The difference in phrasing patterns of Hebrew and Standard Arabic may result in differences in attachment preferences for the ambiguous RC construction. The tendency in Hebrew for relatively short phrases, and the tendency in Standard Arabic towards long phrases, may be translated into differences in the prosodic phrasing of a similar sentence, which in turn may be shown to affect the interpretation of the ambiguous sentence. It is interesting to note that Standard Arabic uses a CS construction to convey genitive relations, and so a comparison of the Hebrew CS with its Arabic counterpart may also be especially revealing with respect to the unique behavior of the CS, as reported in this dissertation.

APPENDICES

Appendix A: Materials

Appendix A-1: Target items used in the three experiments for this dissertation

This appendix lists the 24 experimental items evaluated in Chapter 4; a subset of the 24 (8 items) was used in the production experiment (reported in Chapter 5) and the full 24 items were used in the listening experiment and the silent reading experiment (reported in Chapters 6 and 7, respectively). See Appendix A-2, A-3 and A-4 for the paradigms used in the production, listening and silent reading experiments, respectively.

Each item has an FS (a) and a CS (b) variant, as well as a short RC and long RC variant. Long RCs are extended versions of short RCs; the added material in the long RS is presented in parentheses.

Within the 24 items, there are 2 item subsets: items 1-12 with an animate complex nominal and items 13-24 with an inanimate complex nominal. In all 24 items, N1, N2 and the RC verb agreed in gender and number.

1a. ha-'eda ti'ara 'et ha-me'avte'ax šel ha-sar še-hegi'a (be-tisat layla)
the-witness described ACC the-guard of the-minister who-arrived (in-flight night)
'The witness described the security guard of the minister who arrived (in a night flight)'

1b. ha-'eda ti'ara 'et me'avte'ax ha-sar še-hegi'a (be-tisat layla)
the-witness described ACC the-guard of the-minister who-arrived (in-flight night)
'The witness described the security guard of the minister who arrived (in a night flight)'

2a. ha-'orxim birxu 'et ha-'axot šel ha-kala še-rokedet (rikudey beten)
the-guests congratulated ACC the-sister of the-bride who-dances (dancing belly)
'The guests congratulated the sister of the bride who dances (belly dancing)'

2b. ha-'orxim birxu 'et 'axot ha-kala še-rokedet (rikudey beten)
the-guests congratulated ACC sister the-bride who-dances (dancing belly)
'The guests congratulated the sister of the bride who dances (belly dancing)'

3a. ha-moda'a hex'isa 'et ha-'ima šel ha-tinoket še-hištaala (ši'ul kaved)
the-note angered ACC the-mother of the-baby who-coughed (cough heavy)
'The note angered the mother of the baby who coughed (a heavy cough)'

3b. ha-moda'a hex'isa 'et 'em ha-tinoket še-hištaala (ši'ul kaved)
the-note angered ACC mother the-baby who-coughed (cough heavy)
'The note angered the mother of the baby who coughed (a heavy cough)'

4a. ha-horim hecdiku 'et ha-madrixa šel ha-ganenet še-'išna (be-zman ha-hafsaka)
the-parents justified ACC the-instructor of the-kindergarten.teacher who-smoked (during the-break)
'The parents justified the instructor of the kindergarten teacher who smoked (during the break)'

4b. ha-horim hecdiku 'et madrixat ha-ganenet še-'išna (be-zman ha-hafsaka)
the-parents justified ACC instructor the-kindergarten.teacher who-smoked (during the-break)
'The parents justified the instructor of the kindergarten teacher who smoked (during the break)'

5a. ha-'itona'y ri'ayen 'et ha-bat šel ha-šofetet še-hit'avra (be-'ikvot ha-srefa)
the-jurnalist interviewed ACC the-daughter of the-judge who-was.blinded (as-consequence the-fire)
'The journalist interviewed the daughter of the judge who was blinded (by the fire)'

5b. ha-'itona'y ri'ayen 'et bat ha-šofetet še-hit'avra (be-'ikvot ha-srefa)
the-jurnalist interviewed ACC daughter the-judge who-was.blinded (as-consequence the-fire)
'The journalist interviewed the daughter of the judge who was blinded (by the fire)'

6a. ha-ti'un het'a 'et ha-sanegor šel ha-ganav še-hitlaxeš (be-mahalax ha-mišpat)
the-argument confused ACC the-advocate of the-thief who-whispered (in-course the-trial)
'The argument confused the advocate of the thief who whispered (in the course of the trial)'

6b. ha-ti'un het'a 'et sanegor ha-ganav še-hitlaxeš (be-mahalax ha-mišpat)
the-argument confused ACC advocate the-thief who-whispered (in-course the-trial)
'The argument confused the advocate of the thief who whispered (in the course of the trial)'

7a. ha-morim hergizu 'et ha-mazkira šel ha-menahelet še-'ixra (be-š'aa va-xeci)
the-teachers annoyed ACC the-secretary of the-principal who was.late (in-hour and-half)
'The teachers annoyed the secretary of the principal who was late (in an hour and a half)'

7b. ha-morim hergizu 'et mazkirat ha-menahelet še-'ixra (be-š'aa va-xeci)
the-teachers annoyed ACC secretary the-principal who was.late (in-hour and-half)
'The teachers annoyed the secretary of the principal who was late (in an hour and a half)'

8a. ha-rakezet xipsa 'et ha-nexda šel ha-cayeret še-naxa ('al ha-safsal)
the-director looked.for ACC the-granddaughter of the-painter who rested (on the-bench)
'The director looked for the granddaughter of the painter who rested (on the bench)'

8b. ha-rakezet xipsa 'et nexda ha-cayeret še-naxa ('al ha-safsal)
the-director looked.for ACC granddaughter the-painter who rested (on the-bench)
'The director looked for the granddaughter of the painter who rested (on the bench)'

9a. ha-mištara kivna 'et ha-nahag šel ha-mefaked še-'azav ('et ha-pluga)
the-police directed ACC the-driver of the-commander who-left (ACC the-squadron)
'The police directed the driver of the commander who left (the squadron)'

9b. ha-mištara kivna 'et nahag ha-mefaked še-'azav ('et ha-pluga)
the-police directed ACC driver the-commander who-left (ACC the-squadron)
'The police directed the driver of the commander who left (the squadron)'

10a. ha-'ohadim he'ericu 'et ha-me'amen šel ha-mit'agref še-paraš (le'axar ha-taxarut)
the-fans admired ACC the-coach of the-wrestler who-retired (after the-fight)
'The fans admired the coach of the wrestler who retired (after the fight)'

10b. ha-'ohadim he'ericu 'et me'amen ha-mit'agref še-paraš (le'axar ha-taxarut)
the-fans admired ACC coach the-wrestler who-retired (after the-fight)
'The fans admired the coach of the wrestler who retired (after the fight)'

11a. ha-yedi'a hefti'a 'et ha-te'oma šel ha-zameret še-nifce'a (be-te'unat draxim)
the-news surprised ACC the-twin of the-singer who-was.injured (in.accident road)
'The news surprised the twin of the singer who was injured (in a car accident)'

11b. ha-yedi'a hefti'a 'et te'omat ha-zameret še-nifce'a (be-te'unat draxim)
the-news surprised ACC twin the-singer who-was.injured (in.accident road)
'The news surprised the twin of the singer who was injured (in a car accident)'

12a. Yonatan pagaš 'et ha-ben šel ha-rofe' še-mit'amel (kol yom ba-park)
Jonathan met ACC the-son of the-doctor who-exercises (every day in.the-park)
'Jonathan met the son of the doctor who exercises (every day in the park)'

12b. Yonatan pagaš 'et ben ha-rofe' še-mit'amel (kol yom ba-park)
Jonathan met ACC son the-doctor who-exercises (every day in.the-park)
'Jonathan met the son of the doctor who exercises (every day in the park)'

13a. ha-texnay herim 'et ha-xut šel ha-telefon še-hitnatek (lelo siba)
the-technitian lifted ACC the-cord of the-phone that-disconnected (with.no reason)
'The technician lifted the cord of the phone that disconnected (with no reason)'.

13b. ha-texnay herim 'et xut ha-telefon še-hitnatek (lelo siba)
the-technitian lifted ACC cord the-phone that-disconnected (with.no reason)
'The technician lifted the cord of the phone that disconnected (with no reason)'.

14a. Ruti šatfa 'et ha-xalon šel ha-rexev še-huxlaf (le'axar šanim)
Ruti washed ACC the-window of the-vehicle that-was.replaced (after years)
'Ruti washed the window of the vehicle that was replaced (after years)'

14b. Ruti šatfa 'et xalon ha-rexev še-huxlaf (le'axar šanim)
Ruti washed ACC window the-vehicle that-was.replaced (after years)
'Ruti washed the window of the vehicle that was replaced (after years)'

15a. ha-šomer hestir 'et ha-man'ul šel ha-ša'ar še-nifrac (be-mahalax ha-xufša)
the-guard hid ACC the-lock of the-gate that-was.broken (in-course the-vacation)
'The guard hided the lock of the gate that was broken (during the vacation)'

15b. ha-šomer hestir 'et man'ul ha-ša'ar še-nifrac (be-mahalax ha-xufša)
the-guard hid ACC lock the-gate that-was.broken (in-course the-vacation)
'The guard hid the lock of the gate that was broken (during the vacation)'

16a. ha-'ovdim 'ixsenu 'et ha-'ariza šel ha-madpeset še-nišlexa (mi-mif'al ha-yacran)
the-workers stored ACC the-package of the-printer that-was.sent (from-factory the-manufacture)
'The workers stored the package of the printer that was sent (from the manufacture's factory)'

16b. ha-'ovdim 'ixsenu 'et 'arizat ha-madpeset še-nišlexa (mi-mif'al ha-yacran)
the-workers stored ACC package the-printer that-was.sent (from-factory the-manufacture)
'The workers stored the package of the printer that was sent (from the manufacture's factory)'

17a. ha-yeled šavar 'et ha-beyca šel ha-tarnegolet še-nimce'a (ba-lul 'etmol)
the-boy broke ACC the-egg of the-hen that-was.found (in.the-henhouse yesterday)
'The boy broke the egg of the hen that was found (in the henhouse yesterday)'

17b. ha-yeled šavar 'et beycat ha-tarnegolet še-nimce'a (ba-lul 'etmol)
the-boy broke ACC egg the-hen that-was.found (in.the-henhouse yesterday)
'The boy broke the egg of the hen that was found (in the henhouse yesterday)'

18a. ha-nose'a hexzik 'et ha-yadit šel ha-mizvada še-nikre'aa (be-mahalax ha-mišlo'ax)
the-traveler held ACC the-handle of the-suitcase that-was.torn (in-course the-shipment)
'The traveler held the handle of the suitcase that was torn (in the course of the shipment)'

18b. ha-nose'a hexzik 'et yadit ha-mizvada še-nikre'aa (be-mahalax ha-mišlo'ax)
the-traveler held ACC handle the-suitcase that-was.torn (in-course the-shipment)
'The traveler held the handle of the suitcase that was torn (in the course of the shipment)'

19a. ha-musaxnik garar 'et ha-mexonit šel ha-xevra še-huzkera (be-iton ha-boker)
the-mechanic carried ACC the-car of the-firm that-was.mentioned (in-newspaper the-morning)
'The mechanic carried the car of the firm that was mentioned (in the morning newspaper)'

19b. ha-musaxnik garar 'et mexonit ha-xevra še-huzkera (be-iton ha-boker)
the-mechanic carried ACC car the-firm that-was.mentioned (in-newspaper the-morning)
'The mechanic carried the car of the firm that was mentioned (in the morning newspaper)'

20a. širi kanta 'et ha-sefer šel ha-kurs še-pursam (be-katalog ha-'oniversita)
Širi bought ACC the-book of the-course that-was.advertised (in-catalogue the-university)
'Širi bought the book of the course that was advertised (in the university catalogue)'

20b. širi kanta 'et sefer ha-kurs še-pursam (be-katalog ha-'oniversita)
Širi bought ACC book the-course that-was.advertised (in-catalogue the-university)
'Širi bought the book of the course that was advertised (in the university catalogue)'

21a. ha-adrixal 'icev 'et ha-miznon šel ha-muze'on še-niftax (ba-kayic ha-axaron)
the-architect designed ACC the-cafeteria of the-museum that-was.opened (in.the-summer the-last)
'The architect designed the cafeteria of the museum that was opened (last summer)'

21b. ha-adrixal 'icev 'et miznon ha-muze'on še-niftax (ba-kayic ha-axaron)
the-architect designed ACC cafeteria the-museum that-was.opened (in.the-summer the-last)
'The architect designed the cafeteria of the museum that was opened (last summer)'

22a. ha-šef heš'il 'et ha-mixse' šel ha-sir še-hexlid (me-rov šimuš)
the-chef borrowed ACC the-cover of the-pot that-rusted (from-excess use)
'The chef borrowed the cover of the pot that rusted (from excess of use)'

22b. ha-šef heš'il 'et mixse' ha-sir še-hexlid (me-rov šimuš)
the-chef borrowed ACC cover the-pot that-rusted (from-excess use)
'The chef borrowed the cover of the pot that rusted (from excess of use)'

23a. ha-dayar na'al 'et ha-delet šel ha-dira še-nicbe'aa (be-ceva xadaš)
the-tenant locked ACC the-door of the-apartment that-was.painted (in-paint new)
'The tenant locked the door of the apartment that was painted (in a new paint)'

23b. ha-dayar na'al 'et delet ha-dira še-nicbe'aa (be-ceva xadaš)
the-tenant locked ACC door the-apartment that-was.painted (in-paint new)
'The tenant locked the door of the apartment that was painted (in a new paint)'

24a. ha-po'el 'araz 'et ha-tmuna šel ha-mizraka še-culma (be-magazin ha-'omanut)
the-worker packed ACC the-picture of the-fountain that-was.photoed (in-magazine the art)
'The worker packed the picture of the fountain that was photoed (in the art magazine)'

24b. ha-po'el 'araz 'et tmunat ha-mizraka še-culma (be-magazin ha-'omanut)
the-worker packed ACC picture the-fountain that-was.photoed (in-magazine the art)
'The worker packed the picture of the fountain that was photoed (in the art magazine)'

Appendix A-2: Target items in the production experiment

This appendix shows the paradigm used in the production experiment (reported in Chapter 5). The production experiment used 8 of the 24 experimental items designed for the set of experiments reported in this dissertation; the 8 items used are items 1, 5, 10, 12, 15, 18, 21, 23 from Appendix A-1 above. Each item had an FS and a CS variant, as well as a short RC and a long RC variant. Each item also has a forced-high attachment and a forced-low attachment variant. The paradigm is illustrated in item 1, including the three phrases shown to participants, and the target sentence participants created (which was not shown in the experiment); the remaining items are listed in Appendix A-1 and only the syntactically ambiguous sentence is provided.

1. High attachment: ha-'eda ti'ara 'et (ha-)me'avte'ax (šel) ha-sar
the-witness described ACC (the-)guard (of) the-minister
'The witness described the security guard of the minister'

'eize me'avte'ax?
which guard?
'Which security guard?'

Ha-me'avte'ax še-hegi'a (be-tisat layla)
the-guard who-arrived (in-flight night)
'The security guard who arrived in a night flight'

Low attachment: ha-'eda ti'ara 'et (ha-)me'avte'ax (šel) ha-sar
the-witness described ACC (the-)guard (of) the-minister
'The witness described the security guard of the minister'

'eize sar?
which minister?
'Which minister?'

Ha-sar še-hegi'a (be-tisat layla)
the-minister who-arrived (in-flight night)
'The minister who arrived in a night flight'

Target (created by participants):

ha-'eda ti'ara 'et (ha-)me'avte'ax (šel) ha-sar še-hegi'a (be-tisat layla)
the-witness described ACC (the-)guard (of) the-minister who-arrived (in-flight night)
'The witness described the security guard of the minister who arrived (in a night flight)'

Appendix A-3: Target items in the listening experiment

This appendix shows the paradigm used in the listening experiment (reported in Chapter 6). In all 24 items used in this experiment, N1, N2 and the RC verb agreed in gender and number. Each item has a CS and an FS variant; items with the FS variant have a Phrasing 0, Phrasing 1 and Phrasing 2 variant; items with the CS variant have a Phrasing 0, Phrasing 2-shortRC and Phrasing 2-long RC variant; the slash signs indicate prosodic break positions. Item 1 contains the whole paradigm, including the question and two answers; the remaining items are listed in Appendix A-1, and only the ambiguous target sentences are provided.

1. FS Phrasing 0:
ha-'eda ti'ara / 'et ha-me'avte'ax šel ha-sar še-hegi'a
the-witness described / ACC the-guard of the-minister who-arrived
'The witness described / the security guard of the minister who arrived'

Phrasing 1:
ha-'eda ti'ara 'et ha-me'avte'ax / šel ha-sar še-hegi'a
the-witness described ACC the-guard /of the-minister who-arrived
'The witness described the security guard / of the minister who arrived'

Phrasing 2:
ha-'eda ti'ara / 'et ha-me'avte'ax šel ha-sar / še-hegi'a
the-witness described / ACC the-guard of the-minister / who-arrived
'The witness described / the security guard of the minister / who arrived'

CS Phrasing 0:
ha-'eda ti'ara / 'et me'avte'ax ha-sar še-hegi'a
the-witness described / ACC guard the-minister who-arrived
'The witness described / the security guard of the minister who arrived'

Phrasing 2-short RC:
ha-'eda ti'ara / 'et me'avte'ax ha-sar / še-hegi'a
the-witness described / ACC guard the-minister / who-arrived
'The witness described / the security guard of the minister / who arrived'

Phrasing 2-long RC:
ha-'eda ti'ara / 'et me'avte'ax ha-sar / še-hegi'a be-tisat layla
the-witness described / ACC guard the-minister / who-arrived in-flight night
'The witness described / the security guard of the minister / who arrived in a night flight'

mi hegi'a?	ha-me'avte'ax	ha-sar
Who arrived?	the security guard	the minister

Appendix A-4: Target items in the silent reading experiment

This appendix shows the paradigm used in the silent reading experiment (reported in Chapter 7). In all 24 items used in this experiment, N1, N2 and the RC verb agreed in gender and number. Each item has a CS and an FS variant, as well as a short RC and a long RC variant, as illustrated below in 1. Item 1 contains the whole paradigm; the remaining items are listed in Appendix A-1, and only the ambiguous target sentences are provided.

1a. ha-'eda ti'ara 'et (ha-)me'avte'ax (šel) ha-sar še-hegi'a (be-tisat layla)
the-witness described ACC (the-)guard (of) the-minister who-arrived (in-flight night)
'The witness described the security guard of the minister who arrived (in a night flight)'

mi hegi'a?	ha-me'avte'ax	ha-sar
Who arrived?	the security guard	the minister

Appendix A-5: Filler items used in the experiments for this dissertation

This appendix lists the filler items used in the experiments reported in Chapters 5, 6 and 7. They are distributed into two types: practice filler and fillers.

PRACTICE FILLERS

Items 1-4 were used in the production experiment; items 1-2 were used at the beginning of the experimental task and were not counted towards participants' score on fillers (these are items 7 and 13 in Appendix A-1). Items 2-4 were used with different variants (as illustrated in Appendix A-2) and were used as 'block protectors' at the beginning and end of a break or the whole list of experimental items (these are items 6 and 14 in Appendix A-1). Items 5-7 were used at the beginning of the perception experiments: for each item, the version used in the listening experiment is shown above the slightly modified versions used in the silent reading experiment (the question was the same for both versions; the correct answer is marked with ✓). The slash sign indicates prosodic break positions used in the listening experiment.

1. ha-morim hergizu 'et ha-mazkira šel ha-menahelet še-'ixra be-š'aa va-xeci
   the-teachers annoyed ACC the-secretary of the-principal who was.late in-hour and-half
   'The teachers annoyed the secretary of the principal who was late (in an hour and a half)'

2. ha-texnay herim 'et ha-xut šel ha-telefon še-hitnatek lelo siba
   the-technitian lifted ACC the-cord of the-phone that-disconnected with.no reason
   'The technician lifted the cord of the phone that disconnected (with no reason)'.

3. ha-ti'un het'a 'et ha-sanegor šel ha-ganav še-hitlaxeš be-mahalax ha-mišpat
   the-argument confused ACC the-advocate of the-thief who-whispered in-course the-trial
   'The argument confused the advocate of the thief who whispered (in the course of the trial)'

4. Ruti šatfa 'et ha-xalon šel ha-rexev še-huxlaf le'axar šanim
   Ruti washed ACC the-window of the-vehicle that-was.replaced after years
   'Ruti washed the window of the vehicle that was replaced (after years)'

5. ha-mexva šel ha-mitlamed / rigša 'et ha-mehandes / še-zaxa be-'ot hictaynut me-ha-'igud
   the-gesture of the-practitioner excited ACC the-architect who-won in-token excellency from-the-association
   'The gesture of the practitioner excited the architect who won a token of excellency from the association'

ha-mexva šel ha-mitlamed rigša 'et ha-mehandes še-zaxa be-'ot hictaynut me-ha-'igud
the-gesture of the-practitioner excited ACC the-architect who-won in-token excellency from-the-association
'The gesture of the practitioner excited the architect who won a token of excellency from the association'

mi hitrageš?	a. ha-mitlamed	b. ha-mehandes
Who was excited?	a. the practitioner	✓b. the architect

6 ha-nof ha-marhiv hipnet /'et ha-sabal šel ha-tarmilay
the-view the-amazing hypnotized ACC the-porter of the backpacker
'The amazing view hypnotized the porter of the backpacker'

ha-nof šel ha-harim ha-mušlagim hipnet 'et ha-sabal 'ito ha-tarmilay tiyel
the-view of the-mountains the-snow.covered hypnotized ACC the-porter with.whom the-backpacker hiked
'The view of the snow covered mountains hypnotized the porter with whom the backpacker hiked'

mi hupnat?	a. ha-sabal	b. ha-tarmilay
Who was hypnotized?	✓a. the porter	b. the adventurer

7 be-tekes ha-sium / hecig ha-manxe la-muzmanim / 'et ha-'oman
in-ceremony the-ending presented the host to the people.invited ACC the-artist
'In the ending ceremony the host presented to the people the artist'

be-tekes ha-sium ha-šnati hecig ha-manxe la-muzmanim 'et ha-'oman ha-mefursam
in-ceremony the-ending the-yearly presented the host to the people.invited ACC the-artist the-famous
'In the yearly ending ceremony the host presented to the people the famous artist'

mi hucag?	a. ha-manxe	b. ha-'oman
Who was presented?	a. the host	✓b. the artist

FILLERS

Items 1-72 were used in the perception studies (the listening experiment and the silent reading experiment). For filler items that were not identical in both experiments, the version used in the listening experiment is shown above the version used in the silent reading experiment (for 5 items, the question and two answers were also modified, as indicated below each version). The slash sign indicates prosodic break positions used in the listening experiment.

Within the 72 fillers, there are three item subsets: items 1-24 with a single noun followed by an RC, items 25-36 with a complex nominal but without an RC, and items 36-72 with subordinate clauses. Table 1 shows the distribution of each items subset into further items subsets.

Table 1: distribution of filler items into items subsets

ITEM SUBSETS				ITEM #
Single N with RC	Object position	Short RC	Subject extracted	1-3
			Object extracted	4-6
		Long RC	Subject extracted	7-9
			Object extracted	10-12
	Subject position	Short RC	Subject extracted	13-15
			Object extracted	16-18
		Long RC	Subject extracted	19-21
			Object extracted	22-24
Complex Nominals (no RC)	Construct State	Subject position	Animate	25
			Inanimate	26
		Object position	Animate	27
			Inanimate	28
	Free State	Subject position	Animate	29
			Inanimate	30
		Object position	Animate	31
			Inanimate	32
Subordinate clauses	Sentential complements with complex N		Construct State	33-34
			Free State	35-36
	Adverbial clauses		‘Before’ clauses	49-54
			‘After’ clauses	55-60
			‘When’ clauses	61-66
			‘If’ clauses	67-72

1. ha-xatan tixnen / lehaskir la-xatuna 'et ha-makhela / še-hitparka
   the-groom planned / to.rent for.the-wedding ACC the-chorus / that-separated
   ‘The groom planned to rent for the wedding the chorus that separated’

   mi hitparka? a. ha-xatuna b. ha-makhela
   What separated? a. the wedding ✓b. the chorus

2. ha-karyan he'enik pras / la-mitmoded še-cole'a
the-broadcaster gave award / to-contestant who-limps
'The broadcaster gave an award to the contestant that limps'

ha-karyan he'enik pras la-mitmoded še-cole'a kalot
the-broadcaster gave award to-contestant who-limps lightly
'The broadcaster gave an award to the contestant that limps lightly'

mi col'e?	a. ha-karyan	b. ha-mitmoded
Who limps?	a. the broadcaster	✓b. the contestant

3. ha-banay madad 'im ha-maxšir / 'et ha-kir še-nisdak / be'ikvot re'idat ha-adama
the-constructor measured with the-instrument / ACC the-wall that-cracked / in-consequence shake the-earth.
'The constructor measured with the instrument the wall that was cracked asa result of the earthquake'

ma nisdak?	a. ha-maxšir	b. ha-kir
what was cracked?	a. the instrument	✓b. the wall

ha-banay madad 'im ha-maxšir 'et ha-kir še-nisdak
the-constructor measured with the-instrument ACC the-wall that-cracked
'The constructor measured with the instrument the wall that was cracked'

ma nimdad?	a. ha-maxšir	b. ha-kir
What was measured?	a. the instrument	✓b. the wall

4. ha-psixolog diber 'im ha-pacyent / še-hamirp'aa heznixa
the-psychologist talked with the-patient / who the-clinic neglected
'The psychologist talked with the patient who the clinic neglected'

mi diber?	a. ha-psixolog	b. ha-pacyent
Who talked?	✓a. the psychologist	b. the patient

5. ha-necigim 'irxu 'et ha-mitnadvim / še-ha-'irgun giyes la-mesiba
the-representatives invited ACC the-volunteers / who-the-organization recruited for.the-party
'The representatives invited the volunteers who the organization recruited for the party'

ha-necigim 'irxu 'et ha-mitnadvim še-ha-'irgun giyes
the-representatives invited ACC the-volunteers who-the-organization recruited
'The representatives invited the volunteers who the organization recruited'

mi guysu?	a. ha-necigim	b. ha-mitnadvim
Who were recruited?	a. the representatives	✓b. the volunteers

6. ha-kvišim hyu pkukim / biglal ha-rexovot / še-ha-'iriya sagra le-ša'atayim
the-roads were jammed / because the-streets / that-the-municipality closed for-two-hours
'The roads were jammed because of the streets that the municipality closed for two hours'

ma nisgeru?	a. ha-kvišim	b. ha-rexovot
What were closed?	a. the roads	✓b. the streets

7. ha-mefake'ax piter / 'et ha-šoter še-rašam doxot tnu'aa mezuyafim
the-inspector fired / ACC the-officer who-wrote reports traffic forged
'The inspector fired the officer who wrote forged traffic tickets'

mi putar?	a. ha-mefake'ax	b. ha-šoter
Who was fired?	a. the inspector	✓b. the officer

8. ha-'axot hergi'aa / 'et ha-yoledet še-huvhela le-xadar ha-miyun
the-nurse clamed / ACC the-woman.in.labor who-was.rushed to-room the-emergency.room
'The nursed clamed the woman in labor who was rushed into the emergency room'

mi nirge'aa?	a. ha-axot	b. ha-yoledet
Who was calmed?	a. the nurse	✓b. the woman in labor

9. ha-profesor hitya'ec / 'im ha-sofer še-na'am bifney ha-kenes
the-professor consulted with the-author who-lectured in-face the-conference
'The professor consulted with the author who lectured in the conference'

mi na'am?	a. ha-profesor	b. ha-sofer
Who lectured?	a. The professor	✓b. The author

10. ha-hanhaga hekpi'aa 'et ha-toxnit / še-yacra mehuma ba-medina
the-leadership froze ACC the-plan / that-created chaos in.the-country
'The leadership stopped the plan that created chaos in the country'

mi hekpi'aa?	a. ha-hanhaga	b. ha-toxnit
Who stopped?	✓a. the leadership	b. the plan

11. ha-mefix hitvake'ax / 'im ha-tasritay še-hit'akeš lehosif scena
the-producer argued / with the-screenwriter who-insisted to.add scene
'The producer argued with the screenwriter who insisted on adding a scene'

mi hit'akeš?	a. ha-mefik	b. ha-tasritay
Who insisted?	a. the producer	✓b. the screenwriter

12. ha-mešavek šel ha-sxora / maxar 'et ha-rešet / še-hitpašta be-raxvey ha-'arec
the-distributor of the manufacture / sold ACC the-chain / that spread in-wides the-country
'The distributor of the manufacture sold the chain that spread all over the country'

ma nimkeraa?	a. ha-sxora	b. ha-rešet
What was sold?	a. the manufacture	✓b. the chain

13 ha-tanur še-hitkalkel /mi-šimuš lakuy /yetafked ke-aron 'ixsun
the-oven that-broke / from-use inadequate / will.function as-cabinet storage
'The oven that broke because of inadequate use will function as a storage cabinet'

.

ha-tanur še-hitkalkel yetafked ke-aron 'ixsun
the-oven that-broke will.function as-cabinet storage
'The oven that broke will function as a storage cabinet'

ma-hitkalkel?	a. ha-tanur	b. ha-'aron
What broke?	✓a. the oven	b. the closet

14. ha-'otek še-hechiv / yušlax la-pax / u-bi-mkomo yiraxeš axer
the-copy that-yellowed / will.be.thrown to.the-basket / and-in-its.place will.be purchased another
'The copy that yellowed will be thrown into the basket and in its place another will be purchased'

ma yiraxeš?	a. ha-'otek	b. ha-pax
What yellowed?	✓a. the copy	b. the basket

15. ha-baxura še-histapra / ecel gdol ha-saparim / domaa la-moxeret ba-makolet
the-girl who-got.haircut / at biggest the-hairdressers / resembles to.the-vendor in.the.supermarket
'The girl that got a haircut at the greatest hairdresser resembles the vendor in the supermarket'

ha-baxura še-histapra domaa me'od la-moxeret ba-makolet
the-girl who-got.haircut resembles a lot to.the-vendor in.the.supermarket
'The girl that got a haircut resembles a lot the vendor in the supermarket'

mi histapra?	a. ha-baxura	b. ha-moxeret
Who got haircut?	✓a. the girl	b. the vendor

16. ha-merkazanit ha-xadaša / še-ha-soxnut hexšira la-tafkid /ne'edra me-ha-maxon
the-receptionist the-new / that-the-store trained for.the-job / was.absent from-the-institute
'The new receptionist that the agency trained for the job was absent from the institute'

ha-merkazanit še-ha-soxnut hexšira ne'edra me-ha-maxon be-šel maxala
the-receptionist who-the-agency trained was.absent from-the-institute because-of sickness
'The receptionist whom the agency trained was absent from the institute due to sickness'

mi hexšira?	a. ha-merkazanit	b. ha-soxnut
Who trained?	a. the receptionist	✓b. the agency

17. ha-kirkas še-ha-va'ad / mimen 'et bo'o / helhiv 'et kol ha-cofim
the-circus that-the-board/ financed ACC its.arrival / excited ACC all the-viewers
'The circus that the board financed excited all the viewers'

ha-kirkas še-ha-va'ad mimen helhiv 'et kol ha-cofim
the-circus that-the-board financed excited all the-viewers
'The circus that the board financed excited all the viewers'

mi mimen?	a. ha-kirkas	b. ha-va'ad
Who financed?	a. the circus	✓b. the board

18. ha-macav še-ha-seret mešakef be-verur / hešpi'a 'al ha-noxexim
the-situation that-the-movie depicts in-clarity / affected on the-people.present
'The situation that the movie clearly depicts affected the people present'

ha-macav še-ha-seret mešakef hešpi'a 'al ha-noxexim
the-situation that-the-movie depicts affected on the-people.present
'The situation that the movie depicts affected the people present'

ma mištakef?	✓a. the situation	b. the movie
What is depicted?		

19. ha-xazay ha-vatik / še-medayek ba-taxazit ha-švu'it / mevale 'im ha-bamay ha-ne'erac
the-weatherman the-veteran / who-is.accurate in.the forecast weekly / hang.out with the-director the-adored
'The veteran weatherman who is accurate in the weekly forecast hangs out with the adored director'

ha-xazay še-medayek ba-taxazit ha-švu'it mevale 'im ha-bamay ha-ne'erac txufot
the-weatherman who-is.accurate in.the forecast weekly hang.out with the-director the-adored often
'The weatherman who is accurate in the weekly forecast often hangs out with the adored director'

mi ne'erac?	a. ha-xazay	b. ha-bamay
who is adored?	a. the weatherman	✓b. the director

20. ha-mexonay še-he'esik 'et ha-xašmelay / hexlit linfoš be-eiropa
the-mechanist who-hired ACC the-electrician / decided to.vacate in-Europe
'The mechanist who hired the electriciam decided to vacate in Europe'

mi nofeš?	a. ha-mexonay	b. ha-xašmelay
Who is vacating?	✓a. tThe mechanist	b. the electrician

21. ha-kašiš še-'iter / 'et ha-mumxe le-refu'aa pnimit / kava tor le-bdika
the-old.man who-located / ACC the-specialist for-medicine internal / scheduled appointment for-examination
'The old man who located the specialist for internal medicine scheduled an appointment for an examination'

mi 'utar?	a. ha-kašiš	b. ha-mumxe
Who was located?	a. the old man	✓b. the specialist

22. ha-hitkahalut / še-nigrema me-ha-meriva ha-milulit / sikrena 'et ha-'ovrim ve-šavim
the-gathering / that-was.caused by.the-argument the-verbal / intrigued ACC the-passing and-returning
'The gathering that the was caused by the verbal argument intrigued the passer byes'

ha-hitkahalut še-garma ha-meriva ha-milulit sikrena 'et ha-'ovrim ve-šavim
the-gathering that-caused the-argument the-verbal intrigued ACC the-passing and-returning
'The gathering that the verbal argument caused intrigued the passer byes'

ma sikrena?	a. ha-hitkahalut	b. ha-meriva
What intrigued?	✓a. The gathering	b. The argument

23. ha-mikre / še-ha-xoker badak be-dakdekanut / makše 'al pitron ha-ta'aluma
the-case that-the-investigator examined in-thoroughness makes.difficult on solution the-mystery
'The case that the investigator examined thoroughly makes the solution of the mystery difficult'

ma kaše?	a. ha-mikre	b. ha-pitaron
What is difficult?	a. the case	✓b. the solution

24. ha-kabcan še-ha-'abir hecil / hitbarer ke-musa 'ahavata / šel ha-na'ara
the-beggar who-the-knight rescued / was.cleared as-carrier her.love / of the-girl
'The beggar whom the knight rescued was the source of the girl's love'

mi ne'ehav?	a. ha-kabcan	b. ha-abir
Who is loved?	✓a. the beggar	b. the knight

25. kol 'ozer tabax / šo'ef lihyot be-kšarim / 'im ha-kacav ha-mešuga / ve-ha-yarkan
every helper cook / aspires to.be in-relations / with the-butcher the-crazy / and-the-greengrocer
'Every helper of a cook aspires to be in relation with the crazy butcher and the greengrocer'

mi mešuga?	a. ha-kacav	b. ha-yarkan
Who is crazy?	✓a. the butcher	b. the greengrocer

26. ripud ha-mošav hitlaxlex / me-ha-glida še-yossi axal / 'im ha-'uga
cover the-seat got.dirty / from-the-ice.cream that-Yossi ate / with the cake
'The cover of the seat got dirty from the ice cream that Yossi ate with the cake'

ma lixlexa?	a. ha-glida	b. ha-'uga
What dirtied?	✓a. the ice cream	b. the cake

27. ha-šutaf laxc 'al ha-askan / lehaxzir 'et hanhalat ha-kvuca
the-partner pressured on the-manager / to.return ACC management the-team
'The partner pressured the manager to bring the management of the team back'

mi hulxac?	a. ha-šutaf	b. ha-askan
Who was pressured?	a. the partner	✓b. the manager

28. sufat xol dromit / 'ilca 'et ha-menaka / letate 'et ha-mirpeset šuv
storm wind south / forced ACC the cleaner / to.sweep ACC the-terrace again
'A southern wind storm forced the cleaner to sweep the terrace again'

ma dromit?	a. ha-sufa	b. ha-mirpeset
What is southern?	✓a. the wind	b. the terrace

ha-boc še-ha-gešem saxaf 'ilec 'et ha-menaka lenakot 'et ricpat ha-mirpeset šuv
the-mud that-the-rain carried forced ACC the cleaner to.clean ACC floor the terrace again
'The mud that the rain carried forced the cleaner to clean the floor of the terrace again'

ma nisxaf?	a. ha-boc	b. ha-gešem
What was carried?	✓a. the mud	b. the rain

29. ha-amargan šel ha-dugman / hitlonen befney ha-calam 'al ha-hitrašlut / šel ha-te'uran
the-agent of the-model / complained in.front.of the-photographer about the-clumsiness/ of the-light.engineer
'The agent of the model complained to the photographer about the carelessness of the light engineer'

mi hitrašel?	a. ha-calam	b. ha-te'uran
Who was careless?	a. the photographer	✓b. the light engineer

30. ha-cilum šel ha-pesel / nitram la-ta'aruxa ha-beynle'umit / al yedey ha-galerya
the-photograph of the-statue / was.contributed to.the-exhibition the-international / by the-gallery
'The photograph of the statue was contributed to the international exhibition by the gallery'

mi taram?	a. ha-ta'aruxa	b. Ha-galerya
Who contributed?	a. the exhibition	✓b. the gallery

31. ha-mankal / hezmin la-pgiša 'im ha-maški'a / rak 'et ha-meturgeman šel ha-diplomat
the-director / invited to.the-meeting with the-investor / only ACC the-translator of the-diplomat
'The director invited to the meeting with the investor only the translator of the diplomat'

mi hezmin?	a. ha-mankal	b. ha-maški'a
Who invited?	✓a. the director	b. the investor

32. moše 'ala / 'im ha-sulam ha-ra'u'a la-gag / kedey lexazek 'et ha-'antena šel ha-televizia
Moses came.up / with the-ladder the-shaky to.the-roof / for to.fasten ACC the-antenna of the television
'Moses camp up with the shaky ladder to the roof in order to fasten the antenna of the television'

ma ra'u'a?	a. ha-sulam	b. ha-gag
What is shaky?	✓a. the ladder	b. the roof

33. 'orex ha-šavu'on ta'an / še-ha-katav kibel 'et ha-meyda / mi-makor lo 'amin
editor the-weekly.magazine claimed / that-the-reporter received ACC the-information / from-source not reliable
'The editor of the weekly magazine claimed that the reporter received the information from an unreliable source'

mi hitkabel?	a. ha-mida	b. ha-makor
What was received?	✓a. the information	b. the source

'orex ha-šavu'on ta'an še-ha-katav kibel 'et ha-meyda me-ha-makor me'uxar miday
editor the-weekly.magazine claimed that-the-reporter received ACC the-information from-the-source late too
'The editor of the weekly magazine claimed that the reporter received the information from the source too late'

mi kibel?	a. ha-katav	b. ha-makor
Who received?	✓a. the reporter	b. the source

34. rakdanit ha-lehaka / hodi'aa še-ba'ala nose'a / 'el ha-xore'ograf
dancer the-company / announced that-her.husband is.traveling / to the-choreographer
'The dancer of the company announced that her husband is traveling to the choreographer'

mi nose'a?	a. ha-ba'al	b. ha-xoreograf
Who is traveling?	✓a. the husband	b. the choreographer

35. ha-me'aperet šel ha-koxevet / mamlica še-carxanim / yištamšu ba-mišxa še-ba-'erka
the-makeup.artist of the.star / recommends that-consumers / will.use in.the-cream that-in.the-packet
'The makeup artist of the star recommends that consumers use the cream in the packet'

ma mumlecet?	a. ha-mišxa	b. ha-'erka
What is recommended?	✓a. the cream	b. the packet

36. ha-'ortoped šel ha-sportay / amar še-ha-mitlamedet / ta'azor letapel ba-pcu'aa
the-orthopedic of the-sportsman / said that-the-trainee / will.help to-care.for in.the-injured
'The orthopedic of the sportsman said that the trainee will help care for the injured woman'

mi ta'azor?	a. ha-mitlamedet	b. ha-pcu'aa
Who will help?	✓a. the trainee	b. the injured woman

37. ha-tayar xašav / še-ha-pakax yadrix 'oto la-'oniversita
the-tourist thought /that-the-inspector will.guide him to the-university
'The tourist thought that the inspector will guide him to the university'

ha-tayar xašav še-ha-pakax yadrix 'oto 'el ha-'ir ha-'atika
the-tourist thought that-the-inspector will.guide him to the-city the-old
'The tourist thought that the inspector will guide him to the old city'

mi xašav?	a. ha-tayar	b. ha-pakax
Who thought?	✓a. the tourist	b. the inspector

38. ha-melcarit hitpal'aa / še-ha-gveret lo roca kinu'ax
the-waiter was.surprised / that-the-lady not want desert
'The waiter was surprised that the lady does not want desert

mi hitpal'aa?	a. ha-melcarit	b. ha-geveret
Who was surprised?	✓a. the waiter	b. the lady

39. ha-davar hevxin še-ha-'iš / meroken 'et teyvat ha-do'ar / ba-bkarim
the mailman noticed that-the-man / empties ACC box the-mail / in.the-mornings
'The mailman noticed that the man empties the mailbox in the mornings'

ha-davar hevxin še-ha-'iš meroken 'et teyvat ha-do'ar be-mo'ed kavu'a
the mailman noticed that-the-man empties ACC box the-mail in-time fixed
'The mailman noticed that the man empties the mailbox in a fixed time'

mi meroken?	a. ha-davar	b. ha-'iš
Who empties?	a. the mailman	✓b. the man

40. ha-so'adim samxu / še-ha-me'arxim heci'u le-kinu'ax / rak kafe'
the-diners were.glad / that-the-hosts suggested for-dessert / only coffee
'The diners were glad that the hosts suggested for dessert only coffee'

ha-so'adim samxu še-ha-me'arxim heci'u be-sof ha-'aruxa kafe'
the-diners were.glad that-the-hosts suggested in-end the-meal coffee
'The diners were glad that the hosts suggested at the end of the meal coffee'

mi heci'u?	a. ha-so'adim	b. ha-me'arxim
Who suggested?	a. the diners	✓b. the hosts

41. ha-tošavim hevinu / še-lamrot ha-tlunot / ha-tayasim lo yešanu / 'et maslul ha-nexita
the-residents understood / that-despite the-complaints / the-pilots not will.change / ACC course the-landing
'The residents understood that despite the complaints the pilots would not change the course of the landing'

mi hevinu?	a. ha-tošavim	b. ha-tayasim
Who understood?	✓a. the residents	b. the pilots

42. ha-šadran he'min še-ha-mimca'im / yešaxne'u 'et ha-mitxaze / lehitvadot
the-broadcaster believed that-the-findings / will.convince ACC the-impersonator / to.confess
'The broadcaster believed that the findings would convince the impersonator to confess'
ha-šadran he'min še-ha-mimca'im yešaxne'u 'et ha-mitxaze lehodot ba-'emet
the-broadcaster believed that-the-findings will.convince ACC the-impersonator to.admit in.the-truth
'The broadcaster believed that the findings would convince the impersonator to admit the truth'

mi yištaxne'a?	a. ha-šadran	b. ha-mitxaze
Who will be convinced?	a. the broadcaster	✓b. the impersonator

43. ha-talmidim kivu še-xavreyhem / yityadedu 'im ha-'orxim
the-students hoped that-their.friends would befriend with the-guests
'The students hoped that their friends would befriend theguests'

ha-talmidim kivu še-xavreyhem yuxlu lehityaded 'im š'ar beyt ha-sefer
the-students hoped that-their.friends would befriend with rest the.school
'The students hoped that their friends would befriend the rest of the school'

mi kivu?	a. ha-talmidim	b. ha-xaverim
Who hoped?	✓a. the students	b. the friends

44. ha-sapak hitnacel bifney ha-soxer / 'al ha-'ikuvim ba-mexes
the-supplier apologized to the-merchant / about the-delays in.the-customs
'The supplier apologized to the merchant about the customs delays'

ha-sapak hitnacel bifney ha-soxer še-lo 'amad ba-zamnim biglal 'ikuvim ba-mexes
the-supplier apologized to the-merchant that-not stand in.the-times because delays in.the-custom
'The supplier apologized to the merchant that he did not keep to schedule because of delays in the custom'

mi hitnacel?	a. ha-sapak	b. ha-soxer
Who apologized?	✓a. the supplier	b. the merchant

45. ha-metargelet hevtixa la-marca / še-tesaye'a le-studentim xalašim
the-assistant promised to.the-lecturer / that-she.will.assist to-students weak
'The assistant promised the lecturer that she would assist weak students'

mi tesaye'a?	a. ha-metargelet	b. ha-marca
Who will.assist?	✓a. the assistant	b. the lecturer

46. ha-xayal cipa / še-ha-xoveš it'akev / bešel tna'ey ha-re'ut ha-gru'im
the-soldier expected / that-the-nurse will.be.detained / because conditions the-viewing the-very.bad
'The soldier expected that the-nurse will be detained because of the very bad viewing conditions'

mi it'akev?	a. ha-xayal	b. ha-xoveš
Who will be detained?	a. the soldier	✓b. the nurse

47. ha-dayagim xašešu še-ha-xovevanim / lo yašlixu 'et ha-dagim ha-ktanim / la-agam
the-fishermen were.afraid that-the-amateurs /not will.throw ACC the-fish the-small / to.the-lake
'The fishermen were afraid that the amateurs will not throw the small fish to the lake'

ha-dayagim xašešu še-ha-xovevanim lo izreku 'et ha-dagim ha-ktanim xazara la-agam
the-fishermen were.afraid that-the-amateurs not will.throw.away ACC the-fish the-small back to.the-lake
'The fishermen were afraid that the amateurs will not throw the small fish away back to the lake'

mi xašešu?	a. ha-dayagim	b. ha-xovevanim
Who were afraid?	✓a. the fishermen	b. the amateur

48. mevaker ha-sifrut hoda / še-ha-'axzava me-ha-katava / garma lo lixtov bikoret xarifa
critic the-literature admitted / that-the-disappointment from-the-article / caused him to.write review harsh
'The literature critic admitted that the disappointment from the article caused him to write a harsh review'

mevaker ha-sifrut masar be-tguva še-ha-'axzava me-ha-katava garma lo lixtov bikoret xarifa
critic the-literature said in.response that-the-disappointment from-the-article caused him to.write review harsh
'The literature critic said in response that the disappointment from the article caused him to write a harsh review'

ma 'ixzeva?	a. ha-katava	b. ha-bikoret
What disappointed?	✓a. the article	b. the review

49. lifney še-dani / biker ba-musax, / mixa'el tiken 'et ha-takala
before that-Dani / visited in.the-garage,/ Michael fixed ACC the-failure
'Before Dani visited the garage, Michael fixed the failure'

lifney še-dani biker ba-musax, mixa'el hespik letaken 'et ha-takala
before that-Dani visited in.the-garage, Mixa'el managed to.fix ACC the-failure
'Before Dani visited the garage, Michael managed to fix the failure'

mi tiken?	a. Dani	b. Mixa'el
Who fixed?	a. Dani	✓b. Michael

50. lifney še-ha-atlet / hetxil lehit'amen,/ ha-acan gila xor / ba-na'al ha-yemanit
before that- the-athlete / started to.practice,/ the-sprinter discovered hole / in.the-shoe the-right
'Before the athlete started practicing, the sprinter discovered a hole in his right shoe'

mi gila?	a. ha-atlet	b. ha-acan
Who discovered?	a. the athlete	✓b. the sprinter

51. lifney še-ha-masax nisraf, / ha-maxšev lo her'a simaney gsisa
before that-the-screen was.burned, / the-computer not showed signs dying
'Before the screen was burned, the computer did not show any sign of dying'

ma nisraf?	a. ha-masax	b. ha-maxšev
What was burned?	✓a. the screen	b. the computer

52. ha-lako'ax bikeš / levatel 'et ha-hazmana / lifney še-hašali'ax hevi 'et ha-'oxel
the-customer asked / to.cancel ACC the-order / before that-the-delivery.guy brought ACC the-food
'The customer asked to cancel the order before the delivery guy brought the food'

mi bitel?	a. ha-lako'ax	b. ha-šali'ax
Who canceled?	✓a. the customer	b. the delivery guy

53. ha-mad'an xakar 'et ha-tofa'aa / lifney še-hacibur hit'anyen ba
the-scientist investigated ACC the-phenomena / before that-the-public was.interested in.it
'The scientist investigated the phenomena before the public was interested in'

ha-mad'an xakar 'et ha-tofa'aa lifney še-hacibur hit'anyen lada'at 'aleyha pratim
the-scientist investigated ACC the-phenomena before that-the-public was.interested to.know on.her details
'The scientist investigated the phenomena before the public was interested to know details about it'

mi hit'anyen?	a. ha-mad'an	b. ha-cibur
Who was interested?	a. the scientist	✓b. the public

54. sfinat ha-yamay / 'agna ba-namal / lifney še-ha-sapan / heflig la-yam
ship the-seaman / docked in.the-port / before the-sailor / sailed to.the sea
'The ship of the seaman docked in the port before the sailor sailed to sea'

mi heflig?	a. ha-yamay	b. ha-sapan
Who sailed?	a. the seaman	✓b. the sailor

55. 'axrey še-ha-radio / šider 'et ha-koncert, / hucaf ha-mador be-mixtavim
after that-the-radio / broadcasted ACC the-concert, / was.flooded the-section in-letters
'After the radio broadcasted the concert, the section was flooded with letters'

mi hucaf?	a. ha-radio	b. ha-mador
Who was flooded?	a. the radio	✓b. the section

56. 'axrey še-ha-saxkanit / šipra 'et bicueyha, / panta 'eleyha soxenet noda'at
after that-the-actress / improved ACC her.performances, / turned to.her agent known
'After the actress improved her performances, a well known agent turned to her'

mi hištapra?	a. ha-saxkanit	b. ha-soxenet
Who improved?	✓a. the actress	b. the agent

57. rak axrey še-ha-ne'arim / 'asfu 'et ha-'ecim la-medura, / ha-nagarim samu lev / še-haserim lahem krašim
only after that-the-boys / collected ACC the-woods for.the-bonfire, / the-carpenters put heart / that-missing to.them boards
'Only after the boys collected the woods for the bonfire, the carpenters notices that they are missing baords'

mi 'asfu?	a. ha-ne'arim	b. ha-nagarim
Who collected?	✓a. the boys	b. the carpenters

58. ha-kablan hetxil ba-bniya / axrey še-xatam 'al xoze/ 'im ha-kone
the-contractor started in.the-building / after that-he.signed on contract / with the-buyer
'The contractor started building the building after he signed a contract with the buyer'

ha-kablan hetxil livnot 'et ha-binyan axrey še-xatam 'al xoze 'im ha-kone
the-contractor started to.build ACC the-building after that-he.signed on contract with the-buyer
'The contractor started building the building after he signed a contract with the buyer'

mi bana?	a. ha-kablan	b. ha-kone
Who built?	✓a. the contractor	b. the buyer

59. ha-tove'a / tafas 'et galgal ha-hacala / 'axrey še-ha-mecil saxa 'elav
the-drowning.man / caught ACC wheel the-rescue / after that-the-lifeguard swam till to-him
'The drowning man caught the rescuing wheel after the lifeguard swam up to him'

mi saxa?	a. ha-tove'a	b. ha-mecil
Who swam?	a. the drowning man	✓b. the lifeguard

60. ha-me'alef / heška 'et ha-nemerim ba-kluvim / kše-ha-pilim hetxilu ba-hofa'aa
the-trainer / gave.water ACC the-tigers in.the-cages / when.the-elephants started in.the-show
'The trainer gave water to the tigers after the elephants started the show'

mi šayu?	a. ha-nemerim	b. ha-pilim
Who drank?	✓a. the tigers	b. the elephants

61. lema'ase ha-kabay ne'elam / kvar kše-ha-balaš hitkašer
in.fact the-firefighter disappeared / already when-the-detective called
'In fact the firefighter disappeared already when the detective called'

mi hitkašer?	a. ha-kabay	b. ha-balaš
Who called?	a. the firefighter	✓b. the detective

62. ha-nasi / yaca me-ha-'ulam be-se'ara / kše-ha-yo'ec daxa 'et yešivat ha-boker
the-president / left from-the-hall in-storm / whan-the-adviser postponed ACC meeting the-morning
'The president stormed out of the hall when the advisor postponed the morning meeting'

mi yaca?	a. ha-nasi	b. ha-yo'ec
Who left?	✓a. the president	b. the adviser

63. ha-metapelet nimnema me'at / be-zman še-ha-mišpaxa ši'aše'aa 'et ha-pe'ota
the-nanny napped a.little / in.time that-the-family amused ACC the-baby
'The nanny napped a little while the family amused the baby'

mi nimnema?	a. ha-metapelet	b. ha-pe'ota
Who napped?	✓a. the nanny	b. the baby

64. kše-ha-kosem hexlik / be-'emca ha-mofa ha-marhiv,/ ha-kahal hitpocec mi-cxok
when-the-magician slipped / in-middle the-show the-marvelous, / the- audience exploded of-laughter
'When the magician slipped in the middle of the marvelous show, the audience exploded in laughed'

kše-ha-kosem hexlik be-'emca ha-mofa ha-marhiv, ha-kahal caxak be-mlo grono
when-the-magician slipped in-middle the-show the-marvelous, the- audience laughed in-full throat
'When the magician slipped in the middle of the marvelous show, the audience laugh full heartedly'

mi caxak?	a. ha-kosem	b. ha-kahal
Who laughed?	a. the magician	✓b. the audience

65. kše-ha-politika'I ba le-bikur / be-zman ha-'imun,/ ha-kadurselan heskim lesaxek 'im axyano
when the-politician came to-visit / in.time the-practice,/ the-basketball.player agreed to.play with his-nephew
'When the politician came for a visit during practice, the basketball agreed to play with his nephew'

mi sixek?	a. ha-politika'iy	b. ha-axyan
Who played?	a. the politician	✓b. the nephew

66. kše-ha-taxana acra 'et ha-šidur / be-šel kšayim texnim / ha-misxak ba-migraš hemšix
when-the-station stopped ACC the-broadcast / because-of difficulties technical, / the-game in.the-field continued
'When the station stopped the broadcasting because of technical difficulties, the game in the field continued'

ma ne'ecar?	a. ha-šidur	b. ha-misxak
What was stopped?	✓a. the broadcasting	b. the game

67. ha-moxeret titen / 'et ha-kesef še-ba-kupa / rak 'im ha-šoded / ye'ayem 'al ha-lakoxa
the-saleswoman will.give / the.money that-in.the-register / only if the-robber / will.threaten on the-customer
'the saleswoman will give the money that is in the register only if the robber will threaten the customer'

mi te'uyam?	a. ha-moxeret	b. ha-lakoxa
Who will be threatened?	a. the saleswoman	✓b. the customer

68. 'im ha-sufa timašex be-mlo 'oza, / ha-taxarut tikata
if the-storm will.continue in-full her.strength, / the-competition will.be.stopped
'If the storm will continue in full strength, the competition will be stopped'

'im ha-sufa timašex be-mlo 'oza, ha-taxarut tikata ba-dakot ha-rišonot
if the-storm will.continue in-full her.strength, the-competition will.be.stopped in.the-minutes the-firsts
'If the storm will continue in full strength, the competition will be stopped in the first minutes'

ma timašex?	a. ha-sufa	b. ha-taxarut
What will continue?	✓a. the storm	b. the competition

69. 'im ha-menate'ax ya'avod be-zehirut, / ha-xole yisrod
if the-surgeon will.work in.carefulness, / the-patient will.survive
'If the surgeon will work carefully, the patient will survive'

mi yisrod?	a. ha-menate'ax	b. ha-xole
Who will survive?	a. the surgeon	✓b. the patient

'im ha-menate'ax ya'avod be-zehirut, ha-xole yiša'er ba-xayim
if the-surgeon will.work in.carefulness, the-patient will.stay in.the-life
'If the surgeon will work carefully, the patienti will stay alive'

mi yixye'?	a. ha-menate'ax	b. ha-xole
Who will live?	a. the surgeon	✓b. the patient

70. 'im ha-menace'ax / lo yanxe 'et ha-tizmoret, / ha-kanar yesarev lenagen
if the-conductor not / will.guide ACC the-orchestra, / the-violinist will.refuse to.play
'If the conductor does not guide the orchestra, the violinist will refuse to play'

mi yesarev?	a. ha-menace'ax	b. ha-kanar
Who will refuse?	a. the conductor	✓b. the violinist

71. ha-misrad yiskor 'et ha-me'acev / rak 'im ha-yaxcan / ye'ašer 'et ha-hoca'aa
the-office will.hire ACC the-designer / only if the-PR.person will.approve ACC the-expense
'The office will hire the designer only if the PR person approves the expense'

mi isaxer?	a. ha-me'acev	b. ha-yaxcan
Who will be hired?	✓a. the designer	b. the PR person

72. ha-xadranit / tesader 'et ha-maca'im ha-nekiyim / rak 'im ha-komikayit tit'orer mi-šnata
the-maid / will.arrange ACC the-bed.sheets the-clean / only if the-comedian will.wake from-her.sleep
'The maid will arrange the clean bed sheets only if the comedian wakes up from her sleep'

ha-xadranit tesader 'et ha-maca'im 'im ha-komikayit tit'orer
the-maid will.arrange ACC the-bed.sheets if the-comedian will.wake
'The maid will arrange the bed sheets if the comedian wakes'

mi tit'orer?	a. ha-xadranit	b. ha-komika'it
Who will wake?	a. the maid	✓b. the comedian

Appendix B: Instructions to participants

Appendix B-1: Instructions for the production experiment

Hebrew

הוראות*

בכל אחד מהדפים שבחוברת, יופיעו לפניך: משפט, שאלה הנוגעת למשפט, ותשובה. תוכל לקרוא את שלושתם בשתיקה תחילה, כך שתבין אותם. אחר כך קרא אותם בקול רם, בשטף דיבור אל המיקרופון.

לאחר שקראת את המשפטים ברצף, עליך ליצור משפט חדש על-ידי שילוב המשפטים:
יש לחזור למשפט הראשון ולהמשיכו על-ידי הוספת חלק מהמשפט השני.

עדיף שקודם תגיד לעצמך בראש את המשפט שיצרת, כדי שתדע מה את/ה עומד לומר, ואח"כ תגיד את המשפט שיצרת בקול רם אל המיקרופון. (בדפים הבאים תמצא דוגמות המראות כיצד לשלב את המשפטים).

אנא הקלט כל עמוד בהקלטה נפרדת – באמצע העמוד אין צורך לעצור את ההקלטה (גם אם יש שתיקה בין המשפטים).

השתדל לדבר בצורה שוטפת וטבעית ככל האפשר. המנע מקריאה בסגנון דרמטי או דקלום (כמו מורה או שחקן).
אנו רוצים שתשמע כאילו דיברת אל חברים.

במידה וגמגמת או היססת במהלך הקריאה, לא משנה באיסה חלק של העמוד, יש לחזור על UכלU העמוד מהתחלה.

דוגמה 1:

המשטרה כיוונה את נהג המפקד.
איזה מפקד?
המפקד שעזב את הפלוגה הדרומית.

המשטרה כיוונה את נהג המפקד שעזב את הפלוגה הדרומית.

דוגמה 2:

המורים הרגיזו את המזכירה של המנהלת.
איזו מנהלת?
המנהלת שאיחרה.

המורים הרגיזו את המזכירה של המנהלת שאיחרה.

* לשם הנוחיות, ההוראות כתובות בלשון זכר, אך הן ממוענות לזכר ולנקבה כאחד.

English translation

INSTRUCTION

- In each of the pages in this booklet, you fill find: a sentence, a question relating to this sentence, and an answer. You can read these three in silence first, so you will understand them. Then, read them aloud, fluently, into the microphone.
- After you've read these sentences fluently, go on to create a new sentence by combining these sentences: you need to go back to the first sentence and continue it by combining the last sentence with the first one. (In the next pages you will find examples which show how to combine these sentences).
- Please record every item separately – in the middle of an item you do not need to stop the recording (even if this result sin silence during the recording.

- Try to speak fluently and naturally as much as possible. Avoid reading in a dramatic style or as reciting (as a teacher or an actor).
  We want you to sound as if you are talking to friends.

- **If you stuttered or hesitated during the reading, please repeat the <u>whole</u> page from the beginning.**

* for convenience, the instructions are in written for the male gender, but they are intended for both female and male participants.

EXAMPLE 1: **The police directed the driver of the commander**
Which commander?
The commander who left the southern division.

The police directed the driver of the commander who left the southern division.

EXAMPLE 2: **The teachers annoyed the secretary of the principle**
Which principle?
The principle who was late.

The teachers annoyed the secretary of the principle who was late.

Appendix B-2: Instructions for the listening experiment

Hebrew

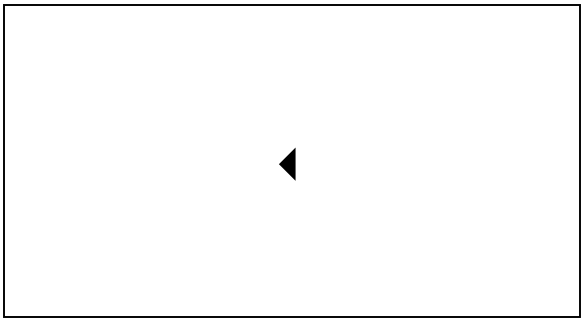
הוראות

מחקר זה בודק כיצד אנשים מפרשים משפטים. משימתך פשוטה מאד: תדרש/י להאזין לסדרה של משפטים. אחרי כל משפט שתשמע/י, תופיע שאלה בנוגע למשמעות המשפט. משימך היא לבחור בתשובה המתאימה מתוך שתי תשובות.

לפניך הוראות מפורטות. הוראות דומות (מקוצרות) יופיעו גם על המסך, בתחילת הניסוי.

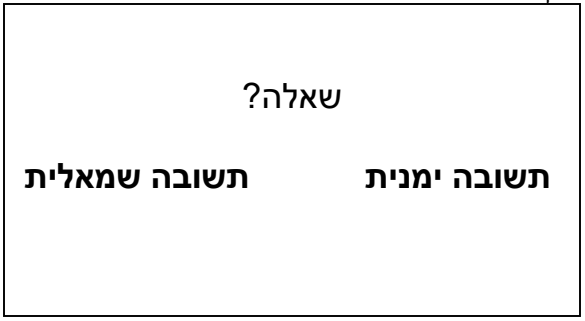
בתחילת כל פריט, יופיע על המסך הסימן: ◂

בכדי להאזין למשפט, לחצ/י על **מקש הרווח**.



תנתן לך אפשרות (אך לא הכרחי) להאזין למשפט פעם אחת נוספת.
להאזנה נוספת, כל שעליך לעשות זה להמתין... (ולא ללחוץ על שום מקש!).
לאחר מספר שניות המשפט יושמע שוב אוטומטית.

אחרי שהאזנת למשפט (פעם או פעמיים), לחצ/י על **מקש השאלה**.
על המסך יופיעו שאלה ושתי תשובות:



לבחירת התשובה **הימנית**,
הקש/י על המקש **הימני** (האדום).

לבחירת התשובה **השמאלית**,
הקש/י על המקש **השמאלי** (הכחול).

ע"פ התשובה שהקשת, תופיע נקודה (אדומה בצד ימין או כחולה בצד שמאל) לסימון תשובתך.

מיד אחר-כך, יופיע פריט חדש על המסך (עם הסימן ◂), ותצטרכ/י ללחוץ על מקש הרווח להאזנה למשפט הבא.

- ינתנו לך 3 משפטי אימון, כדי שתתרגל/י להליך הניסוי. לאחר משפטי האימון ***ולפני שהניסוי יתחיל***, תוכל/י לשאול כל שאלה בנוגע להליך הניסוי.
- בדרך כלל, המשימה אורכת כ-30 דקות.
- הרגע/י ונסה/י להנות. זה איננו מבחן!

אנו מעוניינים באינטואיציות שלך לגבי משמעות המשפטים, לכן יעזור לנו מאד אם תענה/י על השאלות בצורה ספונטנית, מבלי לחשוב יותר מדי.

- ברגע שהתחלת בניסוי, חשוב לנו שיערך ללא הפסקות. אז... אל תשכח/י לכבות את הפלאפון!

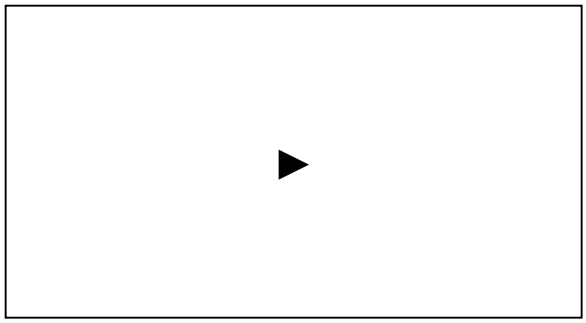
English translation

INSTRUCTION

This study examines how people interpret sentences. Your task is very simple: you will be asked to listen to a list of sentences. After you listen to each sentence, a question relating to the meaning of the sentence will appear. Your task is to choose the appropriate of the two answer-choices.

The following are detailed instructions. Similar (shortened) instructions will appear on the screen as well, at the beginning of the experiment.

At the beginning of each item, the sign: ►
will appear on the screen
To listen to the sentence, click on the **space bar**.

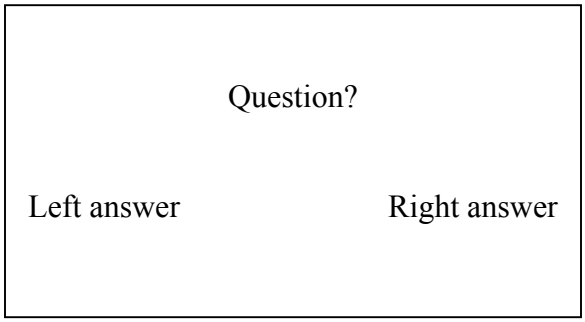


You will have the opportunity (though it's not obligatory) to listen to the sentence one additional time.
For a\that second listening, all you have to do is wait… (do not press on any key!).
After a few seconds the sentence would play again automatically.

After you have listened to the sentence (once or twice), click on the **question key**.
A question with two answer choices will appear on the screen:

To choose the **right** answer,
click on the right key (the red one).

To choose the **left** answer,
Click on the left key (the blue one).



According to the answer you marked, a dot will appear (red on the right side or blue on the left side) to mark your answer.

Immediately after, a new item will appear on the screen (with the mark ►), and you will need to press on the space bar to listen to the next sentence.

- You will be given 3 practice items, so you can practice the procedure. After the practice session ***and before the experiments begins***, you will be able to ask any question regarding the procedure.
- The task usually takes 30 minutes to complete.
- Relax and try to enjoy yourself. This is not a test!
  We are interested in your intuitions about the meaning of the sentences, so it will help us greatly if you answer the questions spontaneously, without thinking too hard.
  Once the experiment starts, it is important that it proceed without interruptions. So… don't forget tot urn off your cell phone!

Appendix B-3: Instructions for the silent reading experiment

Hebrew

הוראות

בתיקייה זו תמצא/י משפטים לקריאה. כל משפט מודפס על דף נפרד. לכל משפט, אנא

תחילה:
קרא/י את המשפט לעצמך בשתיקה, כך שתבינ/י אותו.

אחר כך:
הפוך/הפכי את הדף. בצידו השני של הדף תמצא/י שאלה על משמעות המשפט, עם שתי תשובות. קרא/י את השאלה והתשובות בשתיקה. אחר כך בחר/י את התשובה שנראית לך הגיונית והקף אותה בעיגול.

Uעבור/י על התיקיה, משפט אחד בכל פעם. אל תחזור/י לדף הקודם.

המשימה כולה אורכת כ-35 דקות.

לפני שהניסוי יתחיל, ינתנו לך 3 משפטים עליהם תוכל/י להתאמן. לאחר שתתאמנ/י על משפטים אלו, תוכל/י לשאול כל שאלה שתרצה/י בכל הנוגע למהלך הניסוי.

ברגע שהניסוי יתחיל, אנו מעדיפים שייערך ללא הפסקות. לקראת אמצע הניסוי תגיע לעמוד שיציע לך לקחת הפסקה של מספר דקות.

הרגע/י ונסה/י להנות. זה איננו מבחן.

אנו מעוניינים באינטואיציות שלך לגבי משמעות המשפטים.

יעזור לנו מאד אם תקרא/י בטבעיות, בקצב הנוח לך, ותענה/י על השאלות בצורה ספונטנית מבלי לחשוב יותר מדי.

תוכל/י להתחיל באימון כעת. אנא עצור/י כשתגיע/י לדף המורה לך על הפסקה.

English translation

INSTRUCTIONS

In this folder you will find sentences to read. Each one will be typed on a separate page.

For each sentence, please:

FIRST: read the sentence silently to yourself, so that you understand it.

THEN: Turn the page. You will see a question about the meaning of the sentence, with two answers. Read the question and answers silently. Then select the answer that seems logical to you.

GO THROUGH THE FOLDER ONE SENTENCE AT A TIME. DO NOT GO BACK TO THE PREVIOUS PAGE.

The whole task takes approximately 35 minutes.

Before the experiment begins, there will be 3 sentences for you to practice on. After trying out these sentences, you may ask any questions you have about the procedure

Once the experiment begins, it is best to go through it without interruption. Towards the middle of the booklet you will meet a page offering you to take a few minute's break.

Relax and try to enjoy yourself. This is not a test.

We are interested in your intuitions about what the sentences mean.
It will help us best if you read naturally, at whatever is your comfortable pace, and answer the questions spontaneously without thinking too long or hard.

You may begin the practice session now. Stop when you get to the a page indication a break.

Appendix C: Language history (background) questionnaire

Hebrew

שאלון פרטים אישיים

המידע שיימסר הינו חסוי וישמש לצרכי עבודת המחקר בלבד. השאלון הינו אנונימי.

גיל: ..... מין: ▫ זכר ▫ נקבה

באיזו עיר נולדת? .....

מלבד עברית, באיזו שפה מדברים הוריך בבית? .....

האם עברית היא שפת האם שלך? ..... אם לא, באיזה גיל התחלת ללמוד עברית? .....

באיזה גיל התחלת ללמוד אנגלית? ▫ בערך בגיל 9-10 (כתה ג'-ד')
▫ אם אחר, פרט: .....

מהי רמת הדיבור שלך באנגלית? ▫ שוטף ▫ די טוב ▫ לא כל כך טוב ▫ מעט ▫ בכלל לא

האם את/ה יודע/ת שפות נוספות?

השפה	הגיל בו למדת שפה זו	רמת ידע בשפה (ע"פ הדרוג למעלה)
.....	.....	.....
.....	.....	.....

האם בשלב כלשהו בעברך גרת מחוץ לישראל לתקופה ארוכה משנה?
▫ לא
▫ כן. תאר/י בקצרה היכן, מתי, ולכמה זמן: .....

השכלה (סמנ/י כל מה שרלוונטי. אנא פרט/י את השפה, אם אחרת מעברית, ובאיזו ארץ למדת):

בי"ס יסודי	▫ בעברית	▫ בשפה אחרת.....
בי"ס תיכון	▫ בעברית	▫ בשפה אחרת.....
תואר ראשון	▫ בעברית	▫ בשפה אחרת.....
תואר שני ומעלה	▫ בעברית	▫ בשפה אחרת.....

תעסוקה (אנא סמנ/י רק אחת מהאפשרויות הבאות):

▫ אני מבלה את רוב שעות היום בחברת אנשים (לדוגמה: סטודנט)	▫ אני מבלה את רוב שעות היום עם ניירות (לדוגמה: עו"ד)	▫ אני מבלה את רוב שעות היום עם דברים (לדוגמה: שף)

תודה על השתתפותך!

אנא וודא/י שענית על כל השאלות

English translation

Background questionnaire

All the information that you will provide is confidential and will only be used for purposes of this study. The questionnaire is anonymous.

Age:..... Sex: □ male □ female

City/Country of origin: .....

Besides Hebrew, what language(s) do your parents speak at home? ……

Is Hebrew your native language? ……
If not, how old were you when you started to learn English? …….

How old were you when you started to learn English?
□ Around age 9-10 (3rd –4th grade)
□ If other, please specify ……

How well do you speak English?
□ fluent □ very good □ not so good □ speak very little □ do not speak at all

Do you speak any other language(s)?

Language	Age when learned it	Level (according to the above ranking)
……	……	……
……	……	……

Have you ever lived outside of Israel or a period longer than one year?
□ No
□ Yes. Describe briefly where, when, and for how long:

Education background (check all that apply, and please list the language, if applicable, on the right):

elementary school	□ in Hebrew	□ in another language ……..… ….
high-school	□ in Hebrew	□ in another language …….……..
college	□ in Hebrew	□ in another language ……...
graduate school	□ in English	□ in another language ……...

Occupation (please check only one of the following options):

□ I spend most of my time with people (e.g., a student)	□ I spend most of my time with papers (e.g., a lawyer)	□ I spend most of my time with materials (e.g., a chef)

Thanks for your cooperation!
Please make sure that you have answered all the questions.

Appendix D: Summary tables for Analyses of Variance

The symbols in the rightmost column indicate the significance of the p value, as follows: * = p<.05, ** = p<.01 or p<.005, *** = p<.001. No symbol following the p value indicates p>.1.

Appendix D-1: Production experiment

Perceived Phrasing Data, mean % Phrasing 2

The three factors are: A: Attachment height (low or high)
B: Construction type (FS or CS)
C: RC length (short or long)

Subject-based, % Phrasing 2

SV	SS	DF	MSS	F	p
S	84.11	11	7.65		
A	147.51	1	147.51	21.46	0.001 ***
S*A	75.61	11	6.87		
B	5.51	1	5.51	7.96	0.05 *
S*B	7.61	11	.69		
AB	4.59	1	4.59	9.14	0.05 *
S*AB	5.53	11	.50		
C	128.34	1	128.34	28.08	0.001 ***
S*C	50.28	11	4.57		
AC	82.51	1	82.51	27.41	0.001 ***
S*AC	33.11	11	3.01		
BC	.26	1	.26	.23	0.640
S*BC	12.36	11	1.12		
ABC	.51	1	.51	.69	0.423
SABC	8.11	11	.74		

Item-based, % Phrasing 2

SV	SS	DF	MSS	F	p
S	19.36	7	2.77		
A	221.27	1	221.27	109.78	0.001 ***
S*A	14.11	7	2.02		
B	8.27	1	8.27	27.43	0.005 **
S*B	2.11	7	.30		
AB	6.89	1	6.89	12.11	0.05 *
S*AB	3.98	7	.57		
C	192.52	1	192.52	309.13	0.001 ***
S*C	4.36	7	.62		
AC	123.77	1	123.77	90.16	0.001 ***
S*AC	9.61	7	1.37		
BC	.39	1	.39	2.78	0.140
S*BC	.98	7	.14		
ABC	.77	1	.77	1.30	0.291
SABC	4.11	7	.59		

Subanalyses, Attachment height

The 2 factors are: B: Construction type (FS or CS)
C: RC length (short or long)

Subject-based, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	5.73	11	.52		
B	.02	1	.02	.10	0.754
S*B	2.23	11	.20		
C	2.52	1	2.52	4.12	0.0673
S*C	6.73	11	.61		
BC	.02	1	.02	.10	0.754
S*BC	2.23	11	.20		

Item-based, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	2.97	7	.42		
B	.03	1	.03	.07	0.802
S*B	3.22	7	.46		
C	3.78	1	3.78	10.72	0.05 *
S*C	2.47	7	.35		
BC	.03	1	.03	.07	0.802
S*BC	3.22	7	.46		

Subject-based, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	154.00	11	14.00		
B	10.08	1	10.08	10.16	0.01 *
S*B	10.92	11	.99		
C	208.33	1	208.33	29.89	0.001 ***
S*C	76.67	11	6.97		
BC	.75	1	.75	.45	0.515
S*BC	18.25	11	1.66		

Item-based, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	30.50	7	4.36		
B	15.13	1	15.13	36.83	0.001 ***
S*B	2.88	7	.41		
C	312.50	1	312.50	190.22	0.001 ***
S*C	11.50	7	1.64		
BC	1.13	1	1.13	4.20	0.0796
S*BC	1.88	7	.27		

Sub-Subanalyses, RC length

Factor B: Construction type (FS or CS)

Subject-based, Low attachment, Short RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	203.33	11	18.48		
B	8.17	1	8.17	7.59	0.05 *
S*B	11.83	11	1.08		

Item-based, Low attachment, Short RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	35.00	7	5.00		
B	12.25	1	12.25	22.87	0.005 **
S*B	3.75	7	.54		

Subject-based, Low attachment, Long RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	27.33	11	2.48		
B	2.67	1	2.67	1.69	0.220
S*B	17.33	11	1.58		

Item-based, Low attachment, Long RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	7.00	7	1.00		
B	4.00	1	4.00	28.00	0.005 **
S*B	1.00	7	.14		

Subanalyses, FS

The 2 factors are: A: Attachment height (low or high attachment)
C: RC length (short or long)

Subject-based, FS, % Phrasing 2

SV	SS	DF	MSS	F	p
S	46.00	11	4.18		
A	102.08	1	102.08	22.96	0.001 ***
S*A	48.92	11	4.45		
C	70.08	1	70.08	32.23	0.001 ***
S*C	23.92	11	2.17		
AC	48.00	1	48.00	37.71	0.001 ***
S*AC	14.00	11	1.27		

Item-based, FS, % Phrasing 2

SV	SS	DF	MSS	F	p
S	12.50	7	1.79		
A	153.13	1	153.13	108.54	0.001 ***
S*A	9.88	7	1.41		
C	105.13	1	105.13	841.00	0.001 ***
S*C	.88	7	.13		
AC	72.00	1	72.00	84.00	0.001 ***
S*AC	6.00	7	.86		

Sub-Subanalyses, FS, Attachment height
Factor C: RC length

Subject-based, FS, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	3.46	11	.31		
C	1.04	1	1.04	2.57	0.137
S*C	4.46	11	.41		

Item-based, FS, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	2.44	7	.35		
C	1.56	1	1.56	5.65	0.05 *
S*C	1.94	7	.28		

Subject-based, FS, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	91.46	11	8.31		
C	117.04	1	117.04	38.48	0.001 ***
S*C	33.46	11	3.04		

Item-based, FS, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	19.94	7	2.85		
C	175.56	1	175.56	248.90	0.001 ***
S*C	4.94	7	.71		

Sub-Subanalyses, FS, RC length
Factor A: Attachment height

Subject-based, FS, Short RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	52.46	11	4.77		
A	145.04	1	145.04	36.71	0.001 ***
S*A	43.46	11	3.95		

Item-based, FS, Short RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	9.44	7	1.35		
A	217.56	1	217.56	117.71	0.001 ***
S*A	12.94	7	1.85		

Subject-based, FS, Long RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	17.46	11	1.59		
A	5.04	1	5.04	2.85	0.119
S*A	19.46	11	1.77		

Item-based, FS, Long RC, % Phrasing 2

SV	SS	DF	MSS	F	p
S	3.94	7	.56		
A	7.56	1	7.56	18.02	0.005 **
S*A	2.94	7	.42		

Subanalyses, CS

The 2 factors are: A: Attachment height (low or high attachment)
C: RC length (short or long)

Subject-based, CS, % Phrasing 2

SV	SS	DF	MSS	F	p
S	45.73	11	4.16		
A	50.02	1	50.02	17.07	0.005 **
S*A	32.23	11	2.93		
C	58.52	1	58.52	16.62	0.005 **
S*C	38.73	11	3.52		
AC	35.02	1	35.02	14.15	0.005 ***
S*AC	27.23	11	2.48		

Item-based, CS, % Phrasing 2

SV	SS	DF	MSS	F	p
S	8.97	7	1.28		
A	75.03	1	75.03	63.90	0.001 ***
S*A	8.22	7	1.17		
C	87.78	1	87.78	137.50	0.001 ***
S*C	4.47	7	.64		
AC	52.53	1	52.53	47.64	0.001 ***
S*AC	7.72	7	1.10		

Sub-Subanalyses, CS, Attachment height
Factor C: RC length

Subject-based, CS, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	4.50	11	.41		
C	1.50	1	1.50	3.67	0.0819
S*C	4.50	11	.41		

Item-based, CS, High attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	3.75	7	.54		
C	2.25	1	2.25	4.20	0.0796
S*C	3.75	7	.54		

Subject-based, CS, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	73.46	11	6.68		
C	92.04	1	92.04	16.47	0.005 **
S*C	61.46	11	5.59		

Item-based, CS, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	13.44	7	1.92		
C	138.06	1	138.06	114.54	0.001 ***
S*C	8.44	7	1.21		

Subanalyses, Low attachment

The 2 factors are: B: Construction type (FS or CS)
C: RC length (short or long)

Subject-based, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	154.00	11	14.00		
B	10.08	1	10.08	10.16	0.01 **
S*B	10.92	11	.99		
C	208.33	1	208.33	29.89	0.001 ***
S*C	76.67	11	6.97		
BC	.75	1	.75	.45	0.515
S*BC	18.25	11	1.66		

Item-based, Low attachment, % Phrasing 2

SV	SS	DF	MSS	F	p
S	30.50	7	4.36		
B	15.13	1	15.13	36.83	0.001 ***
S*B	2.88	7	.41		
C	312.50	1	312.50	190.22	0.001 ***
S*C	11.50	7	1.64		
BC	1.13	1	1.13	4.20	0.0796
S*BC	1.88	7	.27		

Sub-Subanalyses, FS, Attachment height, %Phrasing 1
Factor C: RC length

Subject-Based, FS, Low attachment, %Phrasing 1

SV	SS	DF	MSS	F	p
S	91.50	11	8.32		
C	24.00	1	24.00	10.56	0.01 **
S*C	25.00	11	2.27		

Item-Based, FS, Low attachment, %Phrasing 1

SV	SS	DF	MSS	F	p
S	19.75	7	2.82		
C	36.00	1	36.00	42.00	0.001 ***
S*C	6.00	7	.86		

Mean f_0 'rise' of N1 and N2, Elicited productions, RC targets

The three factors are: A: Attachment height (low or high attachment)
B: Construction Type (FS or CS)
C: RC length (short or long)

Mean N1 f_0 'rise'

Subject-based, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	8435.28	11	766.84		
A	518.01	1	518.01	1.11	0.315
S*A	5133.36	11	466.67		
B	1298.01	1	1298.01	14.71	0.005 **
S*B	970.36	11	88.21		
AB	333.76	1	333.76	7.38	0.05 *
S*AB	497.61	11	45.24		
C	219.01	1	219.01	3.74	0.0792
S*C	643.86	11	58.53		
AC	147.51	1	147.51	3.59	0.0848
S*AC	452.36	11	41.12		
BC	61.76	1	61.76	2.85	0.119
S*BC	238.11	11	21.65		
ABC	119.26	1	119.26	3.89	0.0744
SABC	337.61	11	30.69		

Item-based, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	2333.73	7	333.39		
A	310.64	1	310.64	19.86	0.005 **
S*A	109.48	7	15.64		
B	777.02	1	777.02	16.26	0.005 **
S*B	334.61	7	47.80		
AB	221.27	1	221.27	10.37	0.05 *
S*AB	149.36	7	21.34		
C	172.27	1	172.27	13.88	0.01 **
S*C	86.86	7	12.41		
AC	102.52	1	102.52	2.75	0.141
S*AC	260.61	7	37.23		
BC	28.89	1	28.89	3.12	0.120
S*BC	64.73	7	9.25		
ABC	87.89	1	87.89	4.40	0.0741
SABC	139.73	7	19.96		

Mean N2 f_0 'rise'

Subject-based, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	32701.11	11	2972.83		
A	1708.59	1	1708.59	1.06	0.326
S*A	17753.53	11	1613.96		
B	12.76	1	12.76	.12	0.738
S*B	1196.36	11	108.76		
AB	17.51	1	17.51	.34	0.572
S*AB	567.61	11	51.60		
C	3613.76	1	3613.76	15.94	0.005 **
S*C	2493.36	11	226.67		
AC	1881.51	1	1881.51	11.96	0.01 **
S*AC	1730.61	11	157.33		
BC	.84	1	.84	.01	0.928
S*BC	1071.28	11	97.39		
ABC	75.26	1	75.26	2.53	0.140
SABC	327.86	11	29.81		

Item-based, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	5782.00	7	826.00		
A	1122.25	1	1122.25	3.60	0.0996
S*A	2182.25	7	311.75		
B	20.25	1	20.25	.55	0.484
S*B	259.25	7	37.04		
AB	20.25	1	20.25	.37	0.564
S*AB	385.75	7	55.11		
C	2185.56	1	2185.56	12.00	0.0105
S*C	1274.94	7	182.13		
AC	1173.06	1	1173.06	20.69	0.005 **
S*AC	396.94	7	56.71		
BC	5.06	1	5.06	.09	0.777
S*BC	407.94	7	58.28		
ABC	39.06	1	39.06	1.09	0.331
SABC	250.44	7	35.78		

Subanalyses, High attachment data, f_0 'rise'

The 2 factors are: B: Construction Type (FS or CS)
C: RC length (short or long)

Mean N1 f_0 'rise', High attachment data

Subject-based, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	3791.73	11	344.70		
B	157.69	1	157.69	2.29	0.159
S*B	759.06	11	69.01		
C	3.52	1	3.52	.35	0.569
S*C	112.23	11	10.20		
BC	4.69	1	4.69	.26	0.617
S*BC	195.06	11	17.73		

Item-based, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	1251.50	7	178.79		
B	84.50	1	84.50	1.65	0.240
S*B	359.50	7	51.36		
C	4.50	1	4.50	.25	0.636
S*C	128.50	7	18.36		
BC	8.00	1	8.00	.69	0.433
S*BC	81.00	7	11.57		

Mean N2 f_0 'rise', High attachment data

Subject-based, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	25531.42	11	2321.04		
B	30.08	1	30.08	.34	0.573
S*B	979.42	11	89.04		
C	140.08	1	140.08	2.04	0.181
S*C	754.42	11	68.58		
BC	30.08	1	30.08	.43	0.525
S*BC	768.42	11	69.86		

Item-based, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	2312.38	7	330.34		
B	40.50	1	40.50	1.58	0.250
S*B	180.00	7	25.71		
C	78.13	1	78.13	.70	0.431
S*C	784.38	7	112.05		
BC	8.00	1	8.00	.10	0.764
S*BC	575.50	7	82.21		

Sub-Subanalyses, FS f_0 'rise'

The two factors are: A: Attachment Height (high or low)
C: RC length (short or long)

Subject-Based, FS, Mean N1 f_0, 'rise'

SV	SS	DF	MSS	F	p
S	4561.23	11	414.66		
A	841.69	1	841.69	3.31	0.0962
S*A	2798.56	11	254.41		
C	256.69	1	256.69	4.56	0.0559
S*C	618.56	11	56.23		
AC	266.02	1	266.02	4.07	0.0688
S*AC	719.23	11	65.38		

Item-Based, FS, Mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	946.88	7	135.27		
B	528.13	1	528.13	35.94	0.001 ***
S*B	102.88	7	14.70		
C	171.13	1	171.13	10.08	0.05 *
S*C	118.88	7	16.98		
BC	190.13	1	190.13	4.93	0.0618
S*BC	269.88	7	38.55		

Subject-Based, FS, Mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	16020.73	11	1456.43		
B	1036.02	1	1036.02	1.36	0.268
S*B	8382.73	11	762.07		
C	1862.52	1	1862.52	14.47	0.005 **
S*C	1416.23	11	128.75		
BC	1354.69	1	1354.69	17.55	0.005 **
S*BC	849.06	11	77.19		

Item-Based, FS, Mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	3457.38	7	493.91		
B	722.00	1	722.00	3.42	0.107
S*B	1477.50	7	211.07		
C	1200.50	1	1200.50	6.14	0.05 *
S*C	1368.00	7	195.43		
BC	820.13	1	820.13	17.54	0.005 **
S*BC	327.38	7	46.77		

Sub-Subanalyses, FS f_0 'rise', RC length
Factor A: Attachment Height (high or low)

Subject-based, FS, Short RC, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	3593.46	11	326.68		
A	1027.04	1	1027.04	3.98	0.0715
S*A	2841.46	11	258.31		

Item-based, FS, Short RC, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	608.75	7	86.96		
A	676.00	1	676.00	22.01	0.005 **
S*A	215.00	7	30.71		

Subject-based, FS, Long RC, mean N1 f_0

SV	SS	DF	MSS	F	p
S	1586.33	11	144.21		
A	80.67	1	80.67	1.31	0.276
S*A	676.33	11	61.48		

Item-based, FS, Long RC, mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	457.00	7	65.29		
A	42.25	1	42.25	1.87	0.213
S*A	157.75	7	22.54		

Subject-based, FS, Short RC, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	7494.13	11	681.28		
A	2380.04	1	2380.04	5.72	0.05 *
S*A	4573.46	11	415.77		

Item-based, FS, Short RC, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	3212.44	7	458.92		
A	1540.56	1	1540.56	16.12	0.01 **
S*A	668.94	7	95.56		

Subject-based, FS, Long RC, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	9942.83	11	903.89		
A	10.67	1	10.67	.03	0.877
S*A	4658.33	11	423.48		

Item-based, FS, Long RC, mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	1612.94	7	230.42		
A	1.56	1	1.56	.01	0.925
S*A	1135.94	7	162.28		

Subanalyses, CS, f_0 'rise'

The two factors are: A: Attachment Height (high or low)
C: RC length (short or long)

Subject-Based, CS, Mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	4844.42	11	440.40		
A	10.08	1	10.08	.04	0.847
S*A	2832.42	11	257.49		
C	24.08	1	24.08	1.01	0.337
S*C	263.42	11	23.95		
AC	.75	1	.75	.12	0.739
S*AC	70.75	11	6.43		

Item-Based, CS, Mean N1 f_0 'rise'

SV	SS	DF	MSS	F	p
S	1721.47	7	245.92		
A	3.78	1	3.78	.17	0.693
S*A	155.97	7	22.28		
C	30.03	1	30.03	6.43	0.05 *
S*C	32.72	7	4.67		
AC	.28	1	.28	.02	0.906
S*AC	130.47	7	18.64		

Subject-Based, CS, Mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	17876.75	11	1625.16		
A	690.08	1	690.08	.76	0.401
S*A	9938.42	11	903.49		
C	1752.08	1	1752.08	8.97	0.05 *
S*C	2148.42	11	195.31		
AC	602.08	1	602.08	5.48	0.05 *
S*AC	1209.42	11	109.95		

Item-Based, CS, Mean N2 f_0 'rise'

SV	SS	DF	MSS	F	p
S	2583.88	7	369.13		
A	420.50	1	420.50	2.70	0.144
S*A	1090.50	7	155.79		
C	990.13	1	990.13	22.01	0.005 **
S*C	314.88	7	44.98		
AC	392.00	1	392.00	8.57	0.05 *
S*AC	320.00	7	45.71		

Sub-Sunanalyses, CS f_0 'rise', Low attachment
Factor C: RC length (short or long)

Subject-Based, Mean N2 f_0 'rise', CS data, low attachment

SV	SS	DF	MSS	F	p
S	13532.33	11	1230.21		
C	2204.17	1	2204.17	8.85	0.05 *
S*C	2738.83	11	248.98		

Item-Based, Mean N2 f_0 'rise', CS data, low attachment

SV	SS	DF	MSS	F	p
S	2817.44	7	402.49		
C	1314.06	1	1314.06	27.42	0.005 **
S*C	335.44	7	47.92		

Sub-Sunanalyses, CS f_0 'rise', long RC
Factor A: Attachment height (high or low)

Subject-Based, CS, Mean N2 f_0 'rise', long RC

SV	SS	DF	MSS	F	p
S	10267.33	11	933.39		
C	1.50	1	1.50	.00	1.000
S*C	6602.50	11	600.23		

Item-Based, CS, Mean N2 f_0 'rise', long RC

SV	SS	DF	MSS	F	p
S	1106.75	7	158.11		
C	.25	1	.25	.00	1.000
S*C	614.75	7	87.82		

Mean duration of N1 and N2, Elicited productions, RC targets

The three factors are: A: Attachment height (low or high)
B: Construction type (FS or CS)
C: RC length (short or long)

Mean N1 duration

Subject-based, mean N1 duration

SV	SS	DF	MSS	F	p
S	55988.08	11	5089.83		
A	8066.67	1	8066.67	1.92	0.193
S*A	46214.08	11	4201.28		
B	6337.50	1	6337.50	9.63	0.05 *
S*B	7236.25	11	657.84		
AB	6016.67	1	6016.67	8.29	0.05 *
S*AB	7982.58	11	725.69		
C	301.04	1	301.04	1.30	0.278
S*C	2537.71	11	230.70		
AC	3.38	1	3.38	.03	0.868
S*AC	1281.87	11	116.53		
BC	57.04	1	57.04	.36	0.562
S*BC	1752.21	11	159.29		
ABC	15.04	1	15.04	.02	0.885
SABC	7603.71	11	691.25		

Item-based, mean N1 duration

SV	SS	DF	MSS	F	p
S	349154.00	7	49879.14		
A	5476.00	1	5476.00	15.21	0.01 **
S*A	2520.50	7	360.07		
B	19740.25	1	19740.25	3.70	0.0958
S*B	37330.25	7	5332.89		
AB	3906.25	1	3906.25	21.53	0.005 **
S*AB	1269.75	7	181.39		
C	7921.00	1	7921.00	1.73	0.230
S*C	32048.50	7	4578.36		
AC	6.25	1	6.25	.04	0.850
S*AC	1142.75	7	163.25		
BC	4830.25	1	4830.25	.82	0.394
S*BC	40993.75	7	5856.25		
ABC	16.00	1	16.00	.09	0.768
SABC	1186.50	7	169.50		

Mean N2 duration

Subject-based, mean N2 duration

SV	SS	DF	MSS	F	p
S	123136.08	11	11194.19		
A	83308.17	1	83308.17	3.89	0.0742
S*A	235556.08	11	21414.19		
B	2360.17	1	2360.17	1.20	0.297
S*B	21669.58	11	1969.96		
AB	2.67	1	2.67	.00	1.000
S*AB	17364.58	11	1578.60		
C	38400.00	1	38400.00	16.06	0.005 **
S*C	26296.75	11	2390.61		
AC	22693.50	1	22693.50	8.71	0.05 *
S*AC	28648.75	11	2604.43		
BC	2.67	1	2.67	.00	1.000
S*BC	7137.08	11	648.83		
ABC	228.17	1	228.17	.16	0.698
SABC	15825.08	11	1438.64		

Item-based, mean N2 duration

SV	SS	DF	MSS	F	p
S	607712.23	7	86816.03		
A	55755.02	1	55755.02	163.57	0.001 ***
S*A	2386.11	7	340.87		
B	618.77	1	618.77	.18	0.682
S*B	23650.86	7	3378.69		
AB	1.89	1	1.89	.00	1.000
S*AB	5094.73	7	727.82		
C	21352.52	1	21352.52	44.05	0.001 ***
S*C	3393.11	7	484.73		
AC	15159.77	1	15159.77	33.50	0.001 ***
S*AC	3167.86	7	452.55		
BC	178.89	1	178.89	.17	0.693
S*BC	7388.23	7	1055.46		
ABC	159.39	1	159.39	.81	0.397
SABC	1371.73	7	195.96		

Subanalyses, High attachment duration data

The 2 factors are: B: Construction type (FS or CS)
C: RC length (short or long)

Mean N1 duration, High attachment data

Subject-based, mean N1 duration

SV	SS	DF	MSS	F	p
S	60169.92	11	5469.99		
B	2.08	1	2.08	.01	0.938
S*B	3635.92	11	330.54		
C	184.08	1	184.08	1.01	0.336
S*C	2003.92	11	182.17		
AC	6.75	1	6.75	.02	0.885
S*B	3383.25	11	307.57		

Item-based, mean N1 Duration

SV	SS	DF	MSS	F	p
S	156735.50	7	22390.79		
B	3042.00	1	3042.00	1.29	0.294
S*B	16527.50	7	2361.07		
C	4186.13	1	4186.13	1.72	0.231
S*C	17008.38	7	2429.77		
AC	2701.13	1	2701.13	.96	0.359
S*B	19617.38	7	2802.48		

Mean N2 Duration, High attachment data

Subject-based, mean N2 duration

SV	SS	DF	MSS	F	p
S	198760.92	11	18069.17		
B	1102.08	1	1102.08	.64	0.442
S*B	19066.92	11	1733.36		
C	1026.75	1	1026.75	.45	0.515
S*C	24991.25	11	2271.93		
AC	90.75	1	90.75	.07	0.803
S*B	15335.25	11	1394.11		

Item-based, mean N2 duration

SV	SS	DF	MSS	F	p
S	323420.38	7	46202.91		
B	276.13	1	276.13	.10	0.763
S*B	19617.38	7	2802.48		
C	264.50	1	264.50	.84	0.389
S*C	2192.00	7	313.14		
AC	338.00	1	338.00	.80	0.401
S*B	2960.50	7	422.93		

Subanalyses, Low attachment, N1 duration

Sub-Subanalyses, FS data, Low attachment, mean N1 duration
Factor C: RC length (short or long)

Subject-Based, Mean N1 duration, FS data, low attachment

SV	SS	DF	MSS	F	p
S	35607.83	11	3237.08		
C	181.50	1	181.50	.38	0.549
S*C	5226.50	11	475.14		

Item-Based, Mean N1 duration, FS data, Low attachment

SV	SS	DF	MSS	F	p
S	150229.75	7	21461.39		
C	110.25	1	110.25	1.05	0.340
S*C	737.75	7	105.39		

Subanalyses, FS, duration

The two factors are: A: Attachment height (high or low)
C: RC length (short or long)

Subject-Based, Mean N1 duration, FS data

SV	SS	DF	MSS	F	p
S	30704.67	11	2791.33		
A	14008.33	1	14008.33	4.88	0.05 *
S*B	31587.67	11	2871.61		
C	310.08	1	310.08	1.13	0.310
S*C	3011.92	11	273.81		
AC	2.08	1	2.08	.00	1.000
S*BC	5097.92	11	463.45		

Subject-Based, Mean N1 duration, FS data

SV	SS	DF	MSS	F	P
S	266095.88	7	38013.70		
A	9316.13	1	9316.13	20.16	0.005 **
S*B	3234.88	7	462.13		
C	190.13	1	190.13	.76	0.412
S*C	1748.88	7	249.84		
AC	1.13	1	1.13	.01	0.941
S*BC	1321.88	7	188.84		

Subject-Based, Mean N2 duration, FS data

SV	SS	DF	MSS	F	p
S	58572.25	11	5324.75		
A	42126.75	1	42126.75	5.36	0.05 *
S*B	86435.25	11	7857.75		
C	18881.33	1	18881.33	12.12	0.01 **
S*C	17135.67	11	1557.79		
AC	9185.33	1	9185.33	4.25	0.0637
S*BC	23776.67	11	2161.52		

Subject-Based, Mean N2 duration, FS data

SV	SS	DF	MSS	F	P
S	330597.38	7	47228.20		
A	28203.13	1	28203.13	150.55	0.001 ***
S*B	1311.38	7	187.34		
C	12720.13	1	12720.13	16.37	0.005 **
S*C	5440.38	7	777.20		
AC	6105.13	1	6105.13	12.27	0.01 **
S*BC	3482.38	7	497.48		

Sub-Subanalyses, FS duration data, RC length
Factor C: Attachment Height (high or low)

Subject-based, FS, Short RC, mean N1 Duration

SV	SS	DF	MSS	F	p
S	18694.13	11	1699.47		
C	7176.04	1	7176.04	5.18	0.05 *
S*C	15227.46	11	1384.31		

Item-based, FS, Short RC, mean N1 Duration

SV	SS	DF	MSS	F	p
S	126327.00	7	18046.71		
C	4761.00	1	4761.00	38.18	0.001 ***
S*C	873.00	7	124.71		

Subject-based, FS, Long RC, mean N1 Duration

SV	SS	DF	MSS	F	p
S	15022.46	11	1365.68		
C	6834.38	1	6834.38	3.50	0.0881
S*C	21458.13	11	1950.74		

Item-based, FS, Long RC, mean N1 Duration

SV	SS	DF	MSS	F	p
S	141517.75	7	20216.82		
C	4556.25	1	4566.25	8.66	0.05 *
S*C	3683.75	7	526.25		

Subject-based, FS, Short RC, mean N2 Duration

SV	SS	DF	MSS	F	p
S	30092.46	11	2735.68		
C	45327.04	1	45327.04	10.91	0.01 **
S*C	45713.46	11	4155.77		

Item-based, FS, Short RC, mean N2 Duration

SV	SS	DF	MSS	F	p
S	144833.75	7	20690.54		
C	30276.00	1	30276.00	77.01	0.001 ***
S*C	2752.00	7	393.14		

Subject-based, FS, Long RC, mean N2 Duration

SV	SS	DF	MSS	F	p
S	45615.46	11	4146.86		
C	5985.04	1	5985.04	1.02	0.334
S*C	64498.46	11	5863.50		

Item-based, FS, Long RC, mean N2 Duration

SV	SS	DF	MSS	F	p
S	191204.00	7	27314.86		
C	4032.25	1	4032.25	13.82	0.01 **
S*C	2041.75	7	291.68		

Subanalyses, CS duration data

The two factors are: A: Attachment Height (high or low)
C: RC length (short or long)

Subject-Based, Mean N1 Duration, CS data

SV	SS	DF	MSS	F	p
S	32519.67	11	2956.33		
A	75.00	1	75.00	.04	0.852
S*A	22609.00	11	2055.36		
C	48.00	1	48.00	.41	0.534
S*C	1278.00	11	116.18		
AC	16.33	1	16.33	.05	0.832
S*AC	3787.67	11	344.33		

Item-Based, Mean N1 Duration, CS data

SV	SS	DF	MSS	F	p
S	120388.387	7	17198.34		
A	66.13	1	66.13	.83	0.392
S*A	555.38	7	79.34		
C	12561.13	1	12561.13	1.23	0.303
S*C	71293.38	7	10184.77		
AC	21.13	1	21.13	.15	0.713
S*AC	1007.38	7	143.91		

Sub-Subanalyses, CS data, Short RC
Factor A: Attachment height (high or low)

Subject-Based, Mean N2 Duration, CS data, Short RC

SV	SS	DF	MSS	F	p
S	59050.13	11	5368.19		
A	51245.04	1	51245.04	5.55	0.05 *
S*A	101543.46	11	9231.22		

Item-Based, Mean N2 Duration, CS data, Short RC

SV	SS	DF	MSS	F	p
S	131526.94	7	18789.56		
A	34317.56	1	34317.56	67.80	0.001 ***
S*A	3542.94	7	506.13		

Sub-Sunanalyses, CS data, High attachment
Factor C: RC length (short or long)

Subject-Based, Mean N2 Duration, CS data, high attachment

SV	SS	DF	MSS	F	p
S	147880.33	11	13443.67		
C	253.50	1	253.50	.11	0.747
S*C	25573.50	11	2324.		

Item-Based, Mean N2 Duration, CS data, high attachment86

SV	SS	DF	MSS	F	p
S	180640.75	7	25805.82		
C	2.25	1	2.25	.01	0.942
S*C	2760.75	7	394.39		

Appendix D-2: Listening experiment

Percent High attachment, Auditory Questionnaire

Comparison 1: FS (RC=1 word) x Phrasings 0, 1, 2

The 2 factors are: A: Subject/Item Design Groups
C: Phrasing (Phrasing 0, Phrasing 1, Phrasing 2)

Subject-Based, % high attachment

SV	SS	DF	MSS	F	p
A	7338.56	5	1467.71	2.71	0.0537
S	9741.23	18	541.18		
C	22080.66	2	11040.33	68.45	0.001 ***
AC	1896.78	10	189.68	1.18	0.338
S*C	5806.20	36	161.28		

Item-Based, % High attachment

SV	SS	DF	MSS	F	p
A	6521.31	5	1304.26	1.30	0.268
S	114277.79	114	1002.44		
C	111288.14	2	55644.07	131.37	0.001 ***
AC	38487.31	10	3848.73	9.09	0.001 ***
S*C	96572.53	228	423.56		

Sub-analyses, FS, Phrasing

The 2 factors are: A: Subject/Item Design Groups
C: Phrasing (Phrasing 0, Phrasing 1, Phrasing 2)

Suject-Based, % High attachment, Phrasing 0 versus Phrasing 1

SV	SS	DF	MSS	F	p
A	3692.41	5	738.48	2.45	0.0736
S	5430.31	18	301.68		
C	8498.70	1	8498.70	53.74	0.001 ***
AC	1532.32	5	306.46	1.94	0.138
S*C	2846.84	18	158.16		

Item-Based, % High attachment, Phrasing 0 versus Phrasing 1

SV	SS	DF	MSS	F	p
A	16239.63	5	3247.93	4.91	0.001 ***
S	75485.42	114	662.15		
C	43564.68	1	43564.68	124.22	0.001 ***
AC	10071.88	5	2014.38	5.74	0.001 ***
S*C	39980.67	114	350.71		

Subject-Based, % High attachment, Phrasing 0 versus Phrasing 2

SV	SS	DF	MSS	F	p
A	8072.73	5	1614.55	2.47	0.0717
S	11767.74	18	653.76		
C	3006.75	1	3006.75	23.85	0.001 ***
AC	110.28	5	22.06	.17	0.969
S*C	2268.78	18	126.04		

Item-Based, % High attachment, Phrasing 0 versus Phrasing 2

SV	SS	DF	MSS	F	p
A	12906.71	5	2581.34	2.41	0.05 *
S	122140.69	114	1071.41		
C	14643.75	1	14643.75	30.37	0.001 ***
AC	27101.90	5	5420.38	11.24	0.001 ***
S*C	54959.63	114	482.10		

Comparison 2: FS/CS (RC=1 word) x Phrasings 0/2

The 3 factors are: A: Subject/Item Design Groups
B: Construction type (FS or CS)
C: Phrasing (Phrasing 0, Phrasing 2)

Subject-Based, % high attachment

SV	SS	DF	MSS	F	p
A	14559.04	5	2911.81	2.14	0.107
S	24489.57	18	1360.53		
B	8000.80	1	8000.80	45.68	0.001 ***
AB	734.12	5	146.82	.84	0.540
S*B	3152.88	18	175.16		
C	3292.38	1	3292.38	37.42	0.001 ***
AC	170.21	5	34.04	.39	0.851
S*C	1583.83	18	87.99		
BC	406.73	1	406.73	3.41	0.0813
ABC	632.21	5	126.44	1.06	1.06
S*BC	2146.76	18	119.26		

Item-Based, % high attachment

SV	SS	DF	MSS	F	p
A	12309.80	5	2461.96	1.88	0.102
S	148923.13	114	1306.34		
B	39570.64	1	39570.64	67.44	0.001 ***
AB	16274.20	5	3254.84	5.55	0.001 ***
S*B	66890.30	114	586.76		
C	16819.90	1	16819.90	38.90	0.001 ***
AC	25917.51	5	5183.50	11.99	0.001 ***
S*C	49293.85	114	432.40		
BC	1717.63	1	1717.63	3.95	0.05 *
ABC	26400.85	5	5280.17	12.13	0.001 ***
S*BC	49619.17	114	435.26		

Sub-analyses, Construction Type

The 2 factors are: A: Subject/Item Design Groups
C: Phrasing (Phrasing 0, Phrasing 2)

Subject-Based, FS % High attachment

SV	SS	DF	MSS	F	p
A	8072.73	5	1614.55	2.47	0.0717
S	11767.74	18	653.76		
C	3006.75	1	3006.75	23.85	0.001 ***
AC	110.28	5	22.06	.17	0.969
S*C	2268.78	18	126.04		

Item-Based, FS % High attachment

SV	SS	DF	MSS	F	p
A	12906.71	5	2581.34	2.41	0.05 *
S	122140.69	114	1071.41		
C	14643.75	1	14643.75	30.37	0.001 ***
AC	27101.90	5	5420.38	11.24	0.001 ***
S*C	54959.63	144	482.10		

Subject-Based, CS % High attachment

SV	SS	DF	MSS	F	p
A	7220.43	5	1444.09	1.64	0.201
S	15874.71	18	881.93		
C	692.36	1	692.36	8.53	0.01 **
AC	692.14	5	138.43	1.70	0.185
S*C	1461.81	18	81.21		

Item-Based, CS % High attachment

SV	SS	DF	MSS	F	p
A	15677.29	5	3135.46	3.82	0.005 **
S	93672.74	114	821.69		
C	3893.79	1	3893.79	10.10	0.005 **
AC	25216.45	5	5043.29	13.08	0.001 ***
S*C	43953.38	114	385.56		

Comparison 3: CS (Phrasings 2) x Short/long RC

The 2factors are:

A: Subject/Item Design Groups
C: Phrasing (Phrasing 2-short RC/Phrasing 2-long RC)

Subject-Based, % high attachment

SV	SS	DF	MSS	F	p
A	4504.41	5	900.88	.94	0.480
S	17280.17	18	960.01		
C	1.51	1	1.51	.02	0.901
AC	751.57	5	150.31	1.58	0.217
S*C	1716.03	18	95.34		

Item-Based, % high attachment

SV	SS	DF	MSS	F	p
A	4426.50	5	885.30	1.36	0.246
S	74465.71	114	653.21		
C	4.62	1	4.62	.01	0.918
AC	21933.42	5	4386.68	10.06	0.001 ***
S*C	49731.96	114	436.25		

Independent Analyses of Stimuli

Mean Duration of N1 and N2, Item-Based Only

The factor is:
A: 6 Experimental Conditions:
(1) CS-Phrasing 0
(2) CS-Phrasing 2-shortRC
(3) CS-Phrasing 2-long RC
(4) FS-Phrasing 0
(5) FS-Phrasing 1
(6) FS-Phrasing 2

Mean N1 Duration, Stimuli

SV	SS	DF	MSS	F	p
S	1479849.66	23	64341.29		
A	305636.78	5	61127.36	65.40	0.001 ***
S*A	107490.05	115	934.70		

Post-hoc planned comparison: N1, Condition 5 versus average of remaining conditions

SV	SS	DF	MSS	F	p
S	521874.31	23	22690.19		
A	179218.52	1	179218.52	199.24	0.001 ***
S*A	20688.98	23	899.52		

Post-hoc planned comparison: 5 Experimental Conditions (excluding Condition 5)

SV	SS	DF	MSS	F	p
S	1214235.33	23	52792.84		
A	6531.75	4	1632.94	2.06	0.0931
S*A	73095.05	92	794.51		

Mean N2 Duration, Stimuli

SV	SS	DF	MSS	F	p
S	1753317.97	23	76231.22		
A	474357.72	5	94871.54	121.52	0.001 ***
S*A	89783.61	115	780.73		

Post-hoc planned comparison: N2, Average of Conditions 3, 5, 6 versus average of Conditions 1, 2, 4.

SV	SS	DF	MSS	F	p
S	583881.31	23	25386.14		
A	157666.69	1	157666.69	396.07	0.001 ***
S*A	9155.81	23	398.08		

Appendix D-3: Silent reading experiment

Percent High Attachment, Written Questionnaire

The three factors are: A: Subject/Item Design Groups
B: Construction type (FS or CS)
C: RC length (short or long)

Subject-Based, % High attachment

SV	SS	DF	MSS	F	p
A	4873.55	3	1624.52	1.68	0.182
S	54238.35	56	968.54		
B	27093.54	1	27093.54	46.13	0.001 **
AB	4670.67	3	1556.89	2.65	0.0575
S*B	32887.10	56	587.27		
C	9584.79	1	9584.79	37.89	. 0.001 **
AC	2846.32	3	948.77	3.75	0.05 *
S*C	14165.26	56	252.95		
BC	2778.96	1	2778.96	11.97	0.005 **
ASC	4985.06	3	1661.69	7.16	0.001 **
S*BC	13000.18	56	232.15		

Item-Based, % High attachment

SV	SS	DF	MSS	F	p
A	3641.89	3	1213.96	.70	0.566
S	34932.16	20	1746.61		
B	10837.71	1	10837.71	82.25	0.001 **
AB	1783.13	3	594.38	4.51	0.05 *
S*B	2635.33	20	131.77		
C	3833.59	1	3833.59	19.43	0.001 **
AC	675.51	3	225.17	1.14	0.357
S*C	3945.89	20	197.29		
BC	1111.46	1	1111.46	5.43	0.05 *
ASC	849.66	3	283.22	1.38	0.277
S*BC	4094.53	20	204.73		

Sub-analyses, Construction Type

The 2 factors are: A: Subject/Item Design Groups
C: RC Length (short or long)

Subject-Based, FS. % High attachment

SV	SS	DF	MSS	F	p
A	6629.64	3	2209.88	3.13	0.05 *
S	39534.50	56	705.97		
C	11342.85	1	11342.85	61.58	0.001 ***
AC	7509.61	3	2503.20	13.59	0.001 ***
S*C	10314.54	56	184.19		

Item-Based, FS, % High attachment

SV	SS	DF	MSS	F	p
A	4940.70	3	1646.90	2.35	0.10
S	14037.62	20	701.88		
C	4536.71	1	4536.71	27.66	0.001 ***
AC	715.02	3	238.34	1.45	0.25
S*C	3280.73	20	164.04		

Subject-Based, CS, % High attachment

SV	SS	DF	MSS	F	p
A	2914.58	3	971/53	1.14	0.340
S	47590.95	56	849.84		
C	1020.89	1	1020.89	3.39	0.0708
AC	321.77	3	107.26	.36	0.785
S*C	16850.89	56	300.91		

Item-Based, CS, % High attachment

SV	SS	DF	MSS	F	p
A	484.32	3	161.44	.14	0.937
S	23529.87	20	1176.49		
C	408.33	1	408.33	1.72	0.205
AC	810.14	3	270.05	1.13	0.359
S*C	4759.69	20	237.98		

REFERENCES

Abdelghany, H. in preparation. *Prosodic phrasing and modifier attachment in Standard Arabic sentence processing*. Unpublished doctoral dissertation, CUNY Graduate Center, New York, NY.

Abney, S. 1987. The English Noun Phrase in its Sentential Aspect. Unpublished doctoral dissertation, MIT, Cambridge, Mass.

Allbritton, D.W., McKoon, G., & Ratcliff, R. 1996. The reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 714-735.

Amir, N., Almogi, B.-E. & Gal, R. 2004. Perceiving prominence and emotion in speech - a cross lingual study. *Proceedings of Speech Prosody 2004*, Nara, Japan, 375-378.

Amir, N., Silber-Varod, V. & Izre'el, S. 2004. Characteristics on Intonation Unit Boundaries in Spontaneous Spoken Hebrew – Perception and Acoustic Correlates. *Proceedings of Speech Prosody 2004*, Nara, Japan, 677-680.

Anderson, S. R. 1985. Typological distinctions in word formation. In T. Shopen (Ed.) *Language Typology and Syntactic Description*, Vol III *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press.

Argaman, V. & Pearlmutter, N. J. 2002. Lexical semantics as a basis for argument structure frequency biases. In P. Merlo & S. Stevenson (Eds.), *The lexical basis of sentence processing: Formal, computational and experimental perspectives*. Amsterdam: John Benjamins, 303-324.

Augurzky, P. 2005. *Attaching relative clauses in German: The role of implicit and explicit prosody in sentence processing*. Unpublished doctoral dissertation, Universität Leipzig, Germany.

Avineri, Y. 1964. *Yad Ha-Lashon.* Tel Aviv: Yizre'el.

Azar, M. 1976. Analytic and double compounds in daily newspapers. In *Reading in Judaic Studies – Literature, Bible and Linguistics*. Haifa University Press, Haifa (In Hebrew).

Bachenko, J. & Fitzpatrick, E. 1990. A computational grammar of discourse-neutral prosodic phrasing in English. *Computational Linguistics,* 16, 3, 155-170.

Baker, M.C. 1998. Comments on the Paper by Sadock. In S. Lapointe, D. Brenstari & P. Farrel (Eds.) *Morphology and Its Relation to Phonology and Syntax*. Cambridge: Cambridge University Press, 188-212.

Baltazani, M. 2002. The prosodic structure of quantificational sentences in Greek. In M. Andronis, E. Debenport, A. Pycha & K. Yoshimura (Eds.), *Proceedings of the 38th meeting of the Chicago Linguistic Society*, 63-78. CLS.

Bat-El, O. 1993. Parasitic Metrification in the Modern Hebrew Stress System. *The Linguistic Review* 10, 189-210.

Becker, M. 2002. *Hebrew stress: Can you hear those trochees?* Unpublished Masters thesis, Tel Aviv University, Israel.

Becker, M. 2003. Hebrew stress: Can't you hear those trochees? *University of Pennsylvania Working Papers in Linguistics*, 9 (1), 45-57.

Beckman, M. E. & Ayers Elam, G. 1997. Guidelines for ToBI labeling, version 3.0. Manuscript and accompanying speech materials, Ohio State University.

Beckman, M. E. & Hirschberg, J. 1994. The ToBI annotation conventions. Unpublished manuscript, Ohio state University, Columbus, OH.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford, UK: Oxford University Press, 9-54.

Beckman, M. E., & Pierrehumbert, J. B. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.

Benmamoun, E., 2003. Agreement parallelism between sentences and noun phrases: a historical sketch. *Lingua*, 113, 747–764.

Berman, R. 1978. *Modern Hebrew Structure*. Tel Aviv: University Publishing Projects.

Berman, R. 1986. A developmental route: learning about the form and use of complex nominals in Hebrew. *Linguistics*, 25, 1057-1085.

Berman, R. A., 1988. Language knowledge and language use: Binomial constructions in Modern Hebrew. *General Linguistics*, 28, 4, 261-285.

Berman, R. A. & Ravid, D. 1986. Lexicalization of noun compounds. *Hebrew Linguistics*, 24, 5-22 [in Hebrew]

Birch, S. & Clifton, C. Jr. 1995. Focus, Accent and Argument Structure: Effects of Languagte Comprehension. *Language and Speech*, 38, 4, 365-391

Bisetto, A. & Scalise, S. 1999. Compounding: Morphology and/or Syntax? In L. Merev (Ed.) *Boundaries of Morphology and Syntax*. Amsterdam: J. Benjamins, 31-48.

Bolozky, S. 1982. Remarks on Rhythmic Stress in Modern Hebrew. *Journal of Linguistics*, 18, 275-289

Bolozky, S. 2000. Stress placement as a morphological and semantic marker in Israeli Hebrew. *Hebrew Studies*, 41, 53-82.

Borer, H. 1988. On the morphological parallelism between compounds and constructs. In G. Booij & J. van Marle (Eds.). *Yearbook of Morphology*. Dordrecht, Holand: Foris, 45-65.

Borer, H. 1996. The Construct in Review. In J. Lecarme, J. Lowenstamm & U. Shlonsky (Eds.) *Studies in Afroasiatic Grammar*. The Hague: Holland Academic Graphics, 30-61.

Borer, H. 1999. Deconstructing the construct. In K. Johnson & I. Roberts (Eds) *Beyond Principles and Parameters*. Dordrecht: Kluwer, 43-89.

Bradley, D., Fernández, E. M. & Lovric, N. 2003. Overt Prosody in the RC-attachment construction: Elicitation protocols. Poster presented at the 9th Annual Conference on Architectures and Mechanisms for Language Processing, Glasgow, Scotland.

Bradley, D., Fernández, E. M. & Taylor, D. 2003. Prosodic weight versus information load in the relative clause attachment ambiguity. Paper presented at the 16th Annual CUNY Conference on Human Sentence Processing, Cambridge, MA.

Brysbaert, M., & Mitchell, D. C. 1996. Modifier attachment in sentence parsing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology*, 49A, 3, 664-695.

Carlson, K., Clifton, C., Jr., & Frazier, L. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, 45, 58-81

Chayen, M. J. 1973. *The Phonetics of Modern Hebrew*. The Hague: Mouton.

Chomsky, N. 1972. *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1986. Knowledge of Language. New York, Praeger.

Chomsky, N. 1991. Some Notes on Economy of Derivation and Representation. In Robert Freidin (Ed.), *Principles and Parameters in Comparative Grammar*, Cambridge, Mass.: MIT press, 417-454.

Chomsky, N. 1993, A Minimalist Program for Linguistic Theory. In Kenneth Hale and Samuel Jay Keyser, (Eds.), *The View from Building 20*, Cambridge, Mass.: MIT Press, 1-52.

Clifton, C., Jr., Carlson, K., & Frazier, L. 2002. Informative prosodic boundaries. *Language and Speech*, 45, 87-114.

Clifton, C. Jr., Carlson, K. & Frazier, L. 2006. Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic Bulletin and Review*, 13, 5, 854-861.

Cohen, T. 1992. A determiner phrase approach to Hebrew nominals. Unpublished Masters thesus, Bar Ilan University.

Cuetos, F. & Mitchell, D.C. 1988. Cross-linguistic differences in parsing: restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30, 73-105.

Cuetos, F., Mitchell, D.C. & Corley, M.M.B. 1996. Parsing in different languages. In M. Carreiras, J. García-Albea & N. Sebastián-Gallés (Eds.), *Language Processing in Spanih*. Mahwah, N.J.: Lawrence Erlbaum, 145-187.

Danon, G. 2008. Definiteness spreading in the Hebrew construct state. *Lingua*, 118, 872–906.

Desmet, T., Brysbaert, M. & De Baecke, C. 2002. The correspondence between sentence production and corpus frequencies in modifier attachment. *The Quarterly Journal of Experimental Psychology*, 55A, 3, 879–896.

Desmet, T., De Baecke, C., & Drieghe, D. 2005. Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, 21, 4, 453–485.

Dikken, M. den. 1998. Predicate Inversion in DP. In A. Alexiadou and C. Wilder (Eds.) *Possessors, predicates and movement in the Determiner Phrase.* Amsterdam: John Benjamins: 177-214.

D'Imperio, M., Elordieta, G., Frota, S., Prieto, P. & Vigário, M. 2005. Intonational phrasing in Romance: The role of syntactic and prosodic structure. In S. Frota, M. Vigário & M. J. Freitas (Eds.) *Prosodies*. Berlin/New York: Mouton de Gruyter, 59-97.

Dobrovie-Sorin, C. 2000. (In)definiteness Spread: from Romanian Genitives to Hebrew Construct State Nominals. In Motapanyan (Ed), in V. Motapanyane (Ed.), *Comparative Studies in Romanian Syntax*, Oxford: Elsevier, 177-226.

Dobrovie-Sorin, C. 2002. From DPs to NPs: A bare phrase account of genitives. In M. Coene & Y. D'Hulst (Eds.) *From NP to DP. Volume 2: The expression of possession in noun phrases*, Amsterdam/ Philadelphia: John Benjamins.

Doron, E. 1989. Derived nominals in Hebrew. In J. Payne (Ed.), *The Structure of Noun Phrases*. The Hague: Mouton.

Enoch, P. & Kaplan, G. 1969. The physical nature of syllabic stress in Israeli Hebrew. *Leshonenu*, 33, 208-222. [in Hebrew].

Fassi Fehri, A. 1989. Generalized IP structure, Case and VS Word Order. In I. Laka and A. Mahajan (Eds.) *MIT Working Papers in Linguistics*, 10, 75-113.

Fernández, E. M. 2000/2003. Bilingual sentence processing: Relative clause attachment in English and Spanish. Doctoral dissertation, CUNY Graduate Center, New York, NY, 2000. Also in H. Clahsen & L. White (Eds.), *Language Acquisition and Language Disorders, Vol. 29.* Amsterdam: John Benjamins Publishers, 2003.

Fernández, E. M. 2007. How might a rapid serial visual presentation of text affect the prosody projected implicitly during silent reading? *Conferências do V Congresso Internacional da Associação Brasiliera de Lingüística*, 5, 117-154.

Fernández, E. M. & Bradley, D. 2004. Acoustic correlates of phrasing patterns in English and Spanish sentences containing the RC attachment construction. Paper presented at the 5th Annual SUNY/CUNY/NYU Linguistics Mini-Conference, Stony Brook, NY, May 1.

Fernández, E. M., Bradley, D., Igoa, J. M. & Teira, C. 2003. Prosodic phrasing in the RC-attachment ambiguity: Effects of language, RC-length, and position. Paper presented at AMLaP (Architectures and Mechanisms for Languages Processing), Glasgow, Scotland, UK, August 25-27.

Ferreira, F. Henderson, J. M., Anes, M. D., Weeks, P. A. Jr., & McFarlane, D. K. 1996. Effects of Lexical Frequency and Syntactic Complexity in Spoken-Language Comprehension: Evidence From the Auditory Moving-Window Technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 2, 324-335.

Fodor, J.D. 1998. Learning to Parse? *Journal of Psycholinguistic Research*, 27, 2. 285-318.

Fodor, J.D. 2002. Prosodic Disambiguation in Silent Reading. *NELS* 32, 113-132.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 203 – 218.

Fox Tree, J. E., & Meijer, P. J. A. 2000. Untrained speakers' use of prosody in syntactic disambiguation and listeners' interpretations. *Psychological Research/ Psychologische Forschung,* 63, 1-13.

Frazier, L. 1987. On Comprehending Sentences: Syntactic Parsing Strategies. Unpublished doctoral dissertation, University of Connecticut, Connecticut, PA.

Frazier, L. & Clifton, C. Jr. 1996. *Construal*. Cambridge: Massachusetts: MIT Press.

Frazier, L., & Clifton, C., Jr. 1998. Sentence reanalysis, and visibility. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing.*. Amsterdam: Kluwer, 143-176.

Frost R.& Plaut, D. 2005. The word-frequency database for printed Hebrew. word-freq.mscc.huji.ac.il/index.html.

Gee, J. P., & Grosjean, F. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411–458.

Gesenius, W. 1910. *Gesenius' Hebrew Grammar*. Edited by E. Kautzsch, revised by A.E. Cowley. Oxford: Clarendon Press.

Ghini, M. 1993. Ø-formation in Italian: A new proposal. In C. Dyck (Ed.) *Toronto Working Papers in Linguistics*, 23, 2, 41-78. Department of Linguistics, University of Toronto.

Gibson, E., Pearlmutter, N., Canseco-González, E. & Hickock, G. 1996. Recency preference in the human sentence processing mechanism. *Cognition*, 59, 23-59.

Gibson E., Schutze, C.T. & Salomon, A. 1996. The relationship between the frequency and the complexity of linguistic structure, *Journal of Psycholinguistic Rseaech*, 25, 1, 59-92

Gilboy, E., Sopena, J.-M., Clifton, C., & Frazier, L. 1995. Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, 54, 131-167.

Glinert 1989. *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Graf, D. 2000. Stress assignment in the nominal system of Modern Hebrew. In A. Z. Wyner (Ed.), *IATL 7, Proceeding of the Fifteenth Annual Conference of the Israeli Association for Theoretical Lingiustics*, Israel: Bar-Ilan University, 51-70.

Graf, D. & Ussishkin, A. 2002. Emergent iambs: Stress in Modern Hebrew. *Lingua*, 113, 239-270.

Halliday, M.A.K., 1989. *Spoken and Wrtitten Language. Second edition*. Oxford: Oxford University Press.

Halliday M.A.K., 1994. *An Introduction to Functional Grammar. Second edition*. London: Arnold.

Halle, M. & Marantz, A. 1993. Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), The view from building 20: Essays in linguistics in honor of Sylvain Bromberger. Cambridge, Mass.: MIT Press.

Hazout, I. 1991. Verbal Nouns: Theta theoretical studies in Hebrew and Arabic. Unpublished doctoral dissertation, University of Massachusetts, Amherst, MA.

Heller, D. 2002. Possession as a lexical relation: Evidence from the Hebrew Construct State. In L. Mikkelsen and C. Potts (Eds.), *Proceedings of WCCFL 21*. Somerville, MA: Cascadilla Press.

Hellmuth, S. 2004. Prosodic weight and phonological phrasing in Cairene Arabic. *Proceedings of Annual Meeting of Chicago Linguistic Society,* 40.

Hemforth, B. 2006. Crosslinguistic Studies on the Role of Information Structure for Ambiguity Resolution. *Proceedings of the 1st international conference of SFB32: Information structure between linguistic theory and empirical methods.* University of Potsdam, Humboldt, June 6-8, 2006, 17-21.

Hemforth, B. Fernandez, S., Clifton, C. Jr., Frazier, L., Konieczny, L. and Walter, M. 2002. Relative clause attachment in German, English and Spanish: Effects of Position and Length. Unpublished manuscript.

Hemforth, B., Konieczny, L. & Sheepers, C. 2000a. Modifier Attachment: Relative Clauses and Coordinations. In B. Hemforth & L. Konieczny (Eds.), *German Sentence Processing*, Dodrecht: Kluwer Academic Publishers, 159-183.

Hemforth, B., Konieczny, L. & Sheepers, C. 2000b. Syntactic Attachment and Anaphor Resolution: The two Sides of Relative Clause Attachment. In M. W. Crocker, M. J. Pickering & C. Jr. Clifton (Eds.), *Architectures and Mechanisms for Language Processing*, Cambridge, UK: Cambridge University Press, 259-281.

Hemforth, B., Konieczny, L., Scheepers, C. & Strube, G. 1998. Syntactic ambiguity resolution in German. In D. Hillert (Ed.), *Sentence Processing: A Crosslinguistic Perspective*, vol. 31. San Diego, CA: Academic Press, 293-312.

Hemforth, B., Petrone C., D'Imperio, M., Pynte, J., Colonna, S., & Konieczny, L. (2006). Length effects in PP-attachment: Prosodic balancing or balancing information load. *Proceedings of the 28th Annual Conference of the Cognitive Science Society.* Vancover, Canada.

Hirschberg, J. & Avesani, C. 1997. The role of prosody in disambiguating potentially ambiguous utterances in English and Italian. *Proceedings of the ECSA Tutorial and research workshop on intonation*, Athens, 189-192.

Hwang, H. & Schafer, A. J. 2009. Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguist Research*, 38, 2, 151-75.

Isre'el, I. 2005. Intonation Units and the Structure of Spontaneous Spoken Language: A View from Hebrew. In C. Auran, R. Bertrand, C. Chanet, A. Colas, A. Di Cristo, C. Portes, A. Reynier & M. Vion (Eds.) *Proceedings of the IDP05 International Symposium on Discourse-Prosody Interfaces*. Cd ROM and http://aune.lpl.univ-aix.fr/~prodige/idp05/actes/izreel.pdf.

Jones, D. 1956. *The pronunciation of English*. Cambridge university press (first edition 1909)

Jones, D. 1964. *An outline of English phonetics*. Heffer, Cambridge (first edition 1918).

Jun. S-A. 2003. Prosodic Phrasing and Attachment Preferences. *Journal of Psycholinguistic Research*, 32, 2.

Jun, S.-A. 2005. Prosodic Typology in S.-A. Jun (Ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing.* Oxford, UK: Oxford University Press, 430-458.

Jun, S.-A. and Kim, S. 2004 Default phrasing and attachment preference in Korean, *Proceedings of the INTERSPEECH-ICSLP (International Conference on Spoken Language Processing)*, Jeju, Korea.

Jun, S.-A. & .Koike C. 2003 Prosody and Attachment Preference in Japanese: Production and Perception, *Proceedings of the International Workshop on Prosodic Interfaces* . Nantes, France, March 27-29, 2003.

Jun, S.-A., Vicenik, C. & Lofstedt, I. 2007. Intonational Phonology of Georgian. *Poster presented at the Workshop on Intonational Phonology: Understudied or Fieldwork Languages*. A satellite meeting of ICPhS, August 5, 2007, Saarbruecken, Germany

Kamide, Y., Mitchell, D.C., Fodor, J.D., & Inoue, A. 1998. Relative clause attachment ambiguity: Further evidence from Japanese. Poster presented at the 11th annual CUNY Conference on Human Sentence Processing, Rutgers University, New Brunswick, NJ, March, 1998.

Kariaeva, N. 2004. Agreement and locality. Unpublished manuscript, Rutgers University.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph Twenty-Five. The MIT Press: Cambridge.

King, J. W., & Kutas, M. 1995. Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 3, 376–395.

King, W. J., & Just, M. A. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580–602.

Koriat, A., Greenberg, S. & Kreiner, H. 2002. The extraction of structure during reading: Evidence from reading prosody. *Memory and Cognition*, 30, 270-280.

Kraljic, T. & Brennan, S. E. 2005. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology,* 50, 194-231.

Ladd, D. R. 1996. *Intonational phonology.* Cambridge Studies in Linguistics 79. Cambridge: Cambridge University Press

Laufer, A. 1987. *Hangana* ('Intonation'). Jerusalem: Institute for Judaic Studies, Hebrew University of Jerusalem. [in Hebrew].

Laufer, A. 1996. Pauses in fluent speech and punctuation. In M. Bar-Asher (Ed.), *Studies in Hebrew and Jewish Languages Presented to Shelomo Morag*. Jerusalem: The Center for Jewish Languages and Literatures, The Hebrew University of Jerusalem and The Bialik Institute, 277-294. [in Hebrew].

Lehiste, I. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 2, 107-121.

Levi, J. N. 1975. *The Syntax and Semantics of Nonpredicating Adjectives in English.* Unpublished doctoral dissertation, University of Chicago, Chicago.

Levi, J. N. 1976. A Semantic Analysis of Hebrew Compound nominals. In P. Cole (Ed.), *Studies in Modern Hebrew Syntax and Semantics: A Transformational-generative approach*. Amsterdam: North-Holland, 9-55.

Longobardi, G. 1996. On the typological unity on Indo European and Semitic Genitive case. In J. Lecarme, J. Lowenstamm & U. Shlonsky (Eds.), *Studies in Afro-Asiatic Grammar*. The Hague: Holland Academic Publishers, 179-214.

Longobardi, G. 2001. The Structure of DPs: Some Principles, Parameters and Problems. In M. Baltin and C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory*. Blackwell.

Lourenço-Gomes, M.C. 2002. The length of the relative clause as a factor for testing the IPH in Brazilian Portuguese: A preliminary study. Unpublished manuscript.

Lovrić, N. 2003. *Implicit prosody in silent reading: Relative clause attachment in Croatian*. Unpublished doctoral dissertation, CUNY Graduate Center, New York, NY.

Maia, M. Lourenço-Gomes, M.C. & Moraes, J. 2004. Prosodic effects on the reading comprehension and the oral production of ambiguous relative clauses and prepositional phrases in Brazilian Portuguese. Poster presented at the 17th Annual CUNY Human Sentence Processing Conference. College Park, Maryland.

Mak, W. M., Vonk W., Schriefers H. 2006. Processing relative clauses in Dutch: When rocks crush hikers. *Journal of Memory & Language,* 54, 466-490.

Marantz, A. 1997. No escape from Syntax: Don't try Morphological analysis in the privacy of your own Lexicon. Upenn Working Papers in Linguistics, 4, 201-225.

Maynell, L. 1999. Effect of pitch accent placement on resolving relative clause ambiguity in English. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York, NY.

McCauley, A. (in prep). A web-based Hebrew word frequency database. http://homepages.inf.ed.ac.uk/v1amccau/

McGinnis, M. 2002. On the systematic aspect of idioms. Linguistic Inquiry, 33, 4, 665-672.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. 1995. Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, 23, 477–494.

Mitchell, D. 1994. Sentence parsing. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego: Academic Press, pp. 375-409.

Mitchell, D.C. & Cuetos, F. 1991. The origins of parsing strategies. In C. Smith (Ed.), *Current Issues in Natural Language Processing*. Center for Cognitive Science, University of Austin, TX, 1-12.

Mitchell, D., Cuetos, F., Corley, M.B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, 24, 469-488.

Mixdorff, H. & Amir, N. 2002. The Prosody of Modern Hebrew – A Quantitative Study. Paper presented at the speech prosody international conference, Aix-en-Provence, France, 11-13 April 2002.

Most, T., Amir, O. & Tobin, Y. 2000. The Hebrew vowel system: Raw and normalized acoustic data. *Language and Speech*, 43, 3, 295-308.

Noyman, Y. 2000. "sugyot be-inyeney hege ve-cura" (Issues related to phonology and morphology) Ben-Gurion University, Unpublished manuscript [in Hebrew].

Ornan, U. 1964. *Hacerufim Hašemaniyim balašon hairit haxadaša*. Unpublished doctoral dissertation, Hebrew University, Jerusalem.

Peretz, Y. 1968. *Ivrit kahalaxa* ('proper Hebrew') Tel Aviv: Joseph Sreberk Publishing house [in Hebrew].

Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation*. Doctoral dissertation, MIT, published 1988 by IULC.

Pike, K. L. 1958. *The intonation of American English.* Ann Arbor: University of Michigan Press.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* , 90, 6, 2956-2970.

Prieto, P. 2005, Syntactic and eurhythmic constraints on phrasing decisions in Catalan, *Studia Linguistica*, 59, 2-3, special issue on 'Boundaries in Intonational Phonology', M. Horne & M. van Oostendorp (Eds.), 194-222.

Prince A. 1975. T*he Phonology and Morphology of Tiberian Hebrew*. Unpublished doctoral dissertation. MIT, Cambridge, Mass.

Prince, A. & Smolensky, P. 1993/2004. *Optimality theory: Constraint interaction in generative grammar*. Technical report, Rutgers University and University of Colorado at Boulder, 1993. Rutgers Optimality Archive 537, 2002. Revised version published by Blackwell, 2004.

Pynte, J. 1996. Lexical control of within-word eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 958-969.

Pynte, J. & Colonna, S. (2000). Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in French. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.). *Reading as a perceptual process*. Oxford: Elsevier, 529-547.

Quinn, D., Abdelghany, H. & Fodor, J.D. 2000. More evidence of implicit prosody in silent reading: French, English and Arabic relative clauses. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, San Diego

Quinn, D., Fernández, E. M., de Almeida, R. G, Bradley, D., & Fodor J. D. 2001. Prosodic phrasing predicts RC attachment in French and English silent reading. Poster presented at the AMLaP Conference (Architectures and Mechanisms for Language and Processing), Saarbrücken, Germany.

Ravid D. & Shlezinger, Y. 1995. Factors in the Selection of Compound-Types in Spoken and Written Hebrew. *Language Sciences*, 17, 2, 147-179.

Ravid, D., Shlesinger, Y. & Sarel, Z. 1986. Compound Structures - Historical Periods and Registers. Ministry of Education, Jerusalem. [In Hebrew]

Reif, J. A. 1968. *Construct State Nominalization in Modern HebrewI,* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Ritter, E. 1986. NSO order in Modern Hebrew. In J. Mcdonough & B. Blunkett (Eds.) *Proceedings of NELS* 17, 521-537.

Ritter, E. 1988. A head-movement approach to construct-state noun phrases. *Linguistics* 26, 909-929.

Ritter, E. 1991. Two functional categories in Noun Phrases: Evidence from Modern Hebrew. *Syntax and Semantics,* 25, 37-62.

Rosen, H. B. 1956. *Haivrit šelánu - dmuta leor šitot habalšanut*, Um Oved, Tel Aviv. [in Herbew]

Rosen, H. B. 1966. *Ivrit tova*, Jerusalem: Kiryat Sepher (rev. ed.) [in Hebrew]

Rosen, H. 1977. *Contemporary Hebrew*. The Hague: Mouton

Sadock, J.M. 1998. On the Autonomy of Compounding Morphology. In S. Lapointe, D. Brenstari & P. Farrel (Eds.) *Morphology and Its Relation to Phonology and Syntax*. Cambridge: Cambridge University Press, 161-187.

Sandalo, F. & Truckenbrodt, H. 2002. Some Notes on Phonological Phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics*, 42, 285–310.

Schafer, A.J. 1997. *Prosodic Parsing: The Role of Prosody in Sentence Comprehension*. Unpublished doctoral dissertation, University of Massachusetts

Schafer, A., Carlson, K., Clifton, C. Jr., and Frazier, L. 2000b. Focus and the Interpretation of Pitch Accent: Disambiguating Embedded Questions. *Language and Speech*, 43, 1, 75-105.

Schafer, A., Carter, J., Clifton, C. Jr. & Frazier, L. 1996. Focus in Relative Clause Construal. *Language and Cognitive Processes*, 11, 1-2, 135–163.

Schafer, A. J., Speer, S. R. & Warren, P. 2003. Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In J. C. Trueswell & M. Tanenhaus (Eds.), *World situated language use: Perspectives on bridging the product and action tradition.* Cambridge: MIT Press, 209-226..

Schafer, A. J., Speer, S. R., & Warren, P. 2005. Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. K. Tanenhaus & J. Trueswell (Eds.), *World situated language use: Psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition*. Cambridge, MA: MIT Press

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. 2000a. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29, 2, 169-182.

Schriefers, H., Friederici, A. D., & Kühn, K. 1995. The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, 34, 499–520.

Selkirk, E. 1980. The Role of Prosodic Categories in English Word Stress. *Linguistic Inquiry,* 11, 563-605.

Selkirk, E. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, E. 1995. The Prosodic Structure of Function Words. In K. Demuth & J. Morgan (Eds.), *Signal to Syntax*. Lawrence Erlbaum.

Selkirk, E. 2000. The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody: Theory and Experiment*, Dordrecht: Kluwer Academic, 231-262.

Shaked, A. 2001. Predicate Inversion in Hebrew Complex Noun Phrases: The N of N Construction. Unpublished manuscript, CUNY Graduate Center.

Shaked, A. 2006a. Patterns of prosodic phrasing in disambiguated Hebrew nominals: A Post-to-Times study. Poster presented at the 19th Annual CUNY Conference on Human Sentence Processing, March 23-25, 2006, New York, NY.

Shaked, A. 2006b. Competing syntactic and prosodic constraints in Hebrew prosodic phrasing. Paper presented at the GLOW 2006 workshop on Prosodic Phrasing, April 5, 2006, Barcelona, Spain.

Shaked, A. 2007. Competing syntactic and phonological constraints in Hebrew prosodic phrasing. *The Linguistic Review,* 24, 2-3, 169–199.

Shaked, A., Bradley, D. & Fernández, E. M. 2004. Relative Clause Attachment in Hebrew: Free versus Construct State Nominals. Poster presented at the 10th Annual Conference on Architectures and Mechanisms of Language Processing (AMLaP), Aix en Provence, France, September 16-18, 2004.

Shlonsky, U. 2004. The form of Semitic noun phrases. *Lingua* 114, 1465–1526.

Silber-Varod, V. 2005. Characteristics of Prosodic Unit Boundaries in Spontaneous Spoken Hebrew: Perceptual and Acoustic Analysis Unpublished Masters thesis, Tel-Aviv University, Israel.

Siloni, T. 1996. Hebrew Noun Phrases: Generalized Noun Raising. In A. Belletti & L. Rizzi (Eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*. Oxford: Oxford University Press, 239-267.

Siloni, T. 1997. *Noun Phrases and Nominalizations: The Syntax of DP*, Dordrecht: Kluwer Academic Publishers.

Siloni, T., 2001. Construct states at the PF interface. In: Pica, P., Rooryck, J. (Eds.), Linguistic Variation Yearbook, vol. 1. John Benjamins, Amsterdam, pp. 229–266.

Siloni, T., 2003. Prosodic case checking domain: the case of constructs. In J., Lecarme. (Ed.), *Research in Afroasiatic Grammar II*. John Benjamins, 481–510.

Snedeker, J. & Trueswell, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language,* 48, 103-130.

Sproat, R. 1985. *On Deriving the Lexicon*. Cambridge, MA: MIT Working Papers in Linguistics (Published version of Ph.D. dissertation).

Steinhauer, K., & Friederici, A. D. 2001. Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, 30, 3, 267-295.

Stoyneshka, I. 2007. Prosody and ambiguity resolution: Methodological explorations. Poster presented at the 20th annual CUNY sentence Processing conference, La Jolla, California, March 29-31, 2007.

Ten Hacken, P. 1999. Motivated tests for Compounding. *Acta Linguistica Hafniensia*, 31, 27-58.

Thornton, R., MacDonald, M. C., and Arnold, J.E. 2000. The Concomitant Effects of Phrase Length and Informational Content in Sentence Comprehension. *Journal of Psycholinguistic Research*, 2000, 29, 195-203.

Truckenbrodt, H. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry,* 30, 2, 219-255.

Ussishkin, A. 2000. *The emergence of prosody*. Unpublished doctoral dissertation, U. of California, Santa Cruz.

Watson, D. & Gibson, E. 2004. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, 19, 713-755.

Wijnen, F. 2001. Prosody in visual sentence processing. Paper presented at the Prosody in Processing conference, Utrecht institute of Linguistics OTS, Utrecht, NL.