

Syntax in the Making

The emergence of syntactic units
in Finnish conversation

Marja-Liisa Helasvuo

Syntax in the making

Studies in Discourse and Grammar

Studies in Discourse and Grammar is a monograph series providing a forum for research on grammar as it emerges from and is accounted for by discourse contexts. The assumption underlying the series is that corpora reflecting language as it is actually used are necessary, not only for the verification of grammatical analyses, but also for understanding how the regularities we think of as grammar emerge from communicative needs.

Research in discourse and grammar draws upon both spoken and written corpora, and it is typically, though not necessarily, quantitative. Monographs in the series propose explanations for grammatical regularities in terms of recurrent discourse patterns, which reflect communicative needs, both informational and socio-cultural.

Editors

Sandra A. Thompson
University of California at Santa Barbara
Department of Linguistics
Santa Barbara, CA 93106
USA

Paul J. Hopper
Carnegie Mellon University
Department of English
Pittsburgh, PA 15213
USA

Volume 9

Syntax in the Making: The emergence of syntactic units in Finnish conversation
by Marja-Liisa Helasvuo

Syntax in the Making

The emergence of syntactic units
in Finnish conversation

Marja-Liisa Helasvuo

University of Turku

John Benjamins Publishing Company
Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Helasvuo, Marja-Liisa

Syntax in the making : the emergence of syntactic units in Finnish conversation /

Marja-Liisa Helasvuo.

p. cm. (Studies in Discourse and Grammar, ISSN 0928-8929 ; v. 9)

Includes bibliographical references and index.

1. Finnish language--Syntax. I. Title. II. Series.

PH225.H45 2001

494.5415--dc21

2001037992

ISBN 90 272 2619 9 (Eur.) / 1 55619 394 7 (US) (Hb; alk. paper)

© 2001 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

In memory of my father Aulis Alanko

Table of contents

List of tables and figures	XI
Preface	XIII
CHAPTER 1	
Introduction	1
1.1 Grammatical structures as cognitive schemas	4
1.2 Data	9
1.3 Coding and analysis of the data	11
1.3.1 Noun phrase type	11
1.3.2 Semantic class	12
1.3.3 Activation cost	12
1.3.4 Discourse referentiality	13
1.3.5 Grammatical roles	13
1.3.6 Summary	15
1.4 Representing the data	15
1.5 Organization	17
CHAPTER 2	
Dynamics of the clause	19
2.1 What is a clause?	21
2.1.1 Grammar-to-discourse models	22
2.1.2 What are clauses needed for in discourse?	25
2.1.3 Summary	33
2.2 Coding strategies for syntactic functions	33
2.3 Case marking: General principles	35
2.4 Case marking in Finnish	36
2.4.1 Nominative	40
2.4.2 Accusative	46
2.4.3 Partitive	51

2.4.4	Interim summary: The interplay between the grammatical cases in the coding of syntactic roles	56
2.4.5	The adessive and the other oblique cases	57
2.4.6	Summary of case marking	64
2.5	Agreement	64
2.5.1	Sources of agreement markers	65
2.5.2	Agreement: the system and how it is used	66
2.5.3	Cross-referencing or indexing	73
2.5.4	Summary: Agreement in the coding of grammatical relations	75
2.6	Word order	76
2.6.1	Word order and the preverbal position	76
2.6.2	Word order and type of subject	79
2.6.3	Summary on word order	81
2.7	Summary: The clause revisited	82

CHAPTER 3

Grammaticization of the subject role 85

3.1	Is there a unified subject role in Finnish?	85
3.1.1	Subject and clause type	85
3.1.2	The definition of subject from a discourse perspective	88
3.2	What do subjects do in discourse?	89
3.2.1	Information flow and the subject role	90
3.2.2	Split patterning and the grammaticization of the subject role	94
3.3	Subject and non-subject	97
3.3.1	Existential constructions	97
3.3.2	Subjects and open quantification	101
3.4	Summary on the grammaticization of the subject	103

CHAPTER 4

Free NPs 105

4.1	What is a free NP?	105
4.2	Free NPs as syntactic units: A cross-linguistic perspective	108
4.3	Case marking of free NPs	113

4.4	Discourse functions of free NPs	117
4.4.1	Identifying and classifying free NPs	117
4.4.2	Theme and orientation	123
4.4.3	Topic constructions	126
4.4.4	Other	130
4.5	Summary	131
CHAPTER 5		
	Intonation and syntactic structuring	133
5.1	Intonation and constituent structure	135
5.1.1	The emergence of constituent structure in discourse	135
5.1.2	Intonation and constituency in Finnish	136
5.1.3	Summary	140
5.2	Intonational grouping and clausal organization	140
5.2.1	Oblique NPs and clause structure	141
5.2.2	Intonational grouping and the clause core	145
5.3	Summary	149
CHAPTER 6		
	Conclusion	151
	References	155
	Appendix: Glossing and transcription conventions	167
	Glossing	167
	Nominal markings	167
	Verbal markings	168
	Other	168
	Transcription conventions	168
	Transitional continuity	168
	Accent and lengthening	168
	Quality	169
	Vocal noises	169
	Other	169
	Name index	171
	Subject index	175

List of tables and figures

Table 1.	The Finnish case system	37
Table 2.	Case marking of NPs: discourse frequencies	37
Table 3.	Distribution of nominative NPs across syntactic roles	39
Table 4.	Case marking of the core roles. Personal pronouns	43
Table 5.	Case marking of the core roles. Other singular pronouns and singular full NPs	44
Table 6.	Case marking of the core roles. Plural full NPs and plural pronouns (except for personal pronouns)	44
Table 7.	Object marking in Finnish. Personal pronouns	46
Table 8.	Object marking in Finnish. Full NPs and pronouns other than personal ones	48
Table 9.	The distribution of human referents across cases	54
Table 10.	The distribution of further mentions across cases with which the initial mention was made	55
Table 11.	Agreement system in Finnish. Verbal agreement morphology	67
Table 12.	Word order in clauses with 3rd person pronominal subjects	79
Table 13.	Word order in clauses with 1st or 2nd person subjects	80
Table 14.	Word order in clauses with full NP subjects	80
Table 15.	Type of subject NP across clause type	86
Table 16.	Activation cost and grammatical role	90
Table 17.	Referent tracking and grammatical role	92
Table 18.	Distribution of human referents across grammatical role	92
Figure 1.	Type of subject and grammatical role	87

Preface

This book is a love story. As this may not be immediately obvious, a few words are in order. When I was a young student of linguistics at the University of Helsinki, we had a visit from an eminent linguist, professor Sandra A. Thompson. The topic had to do with subjects and objects, and the goal of the talk was to show why these grammatical roles should not be considered universal. At the discussion session after the talk, professor Thompson claimed, “The reason why we study linguistics is of course the beauty of grammar!” I wrote this down in my note book, thinking how true it was. I love grammar and its intricate characteristics, and I love to explore the ways it is being used. I have written this book to all the people who share this love.

Later on, I received the opportunity to pursue my graduate studies under the supervision of Sandy Thompson and her colleagues at the University of California, Santa Barbara. This book grew out of my Ph.D. thesis, which was completed in 1997 with Sandy Thompson as the committee chair. I feel privileged to have worked with her, for she is the greatest mentor and teacher I have ever known: She has the gift of helping other people say what they think. She has also been a great source of inspiration, not only in our private discussions, but also, through her numerous writings.

My understanding of the emergence of clausal argument relations has been shaped by Jack Du Bois’s thinking. He and Wally Chafe have both read my work and provided insightful comments, for which I am grateful. Wally Chafe has shown me the importance of building an overall framework of how grammar works. I am also grateful to Auli Hakulinen for commenting on my work and providing important references.

For me, the linguistics department at UCSB was an ideal place where an encouraging and supportive atmosphere was combined with an inspiring intellectual environment. I remember one time during Jack and Sandy’s seminar on conversation and grammar when zeros were so hot that after class we all packed into Jack’s tiny office and everybody was eager to express his or her own view on whether one could speak about zero arguments in his or her language. Now I see this discussion continued in the writings of all of us

wherever we are in the world, and it gives me great joy to follow the discussion and take part in it. I am especially grateful to Mike Ewing, Agnes Kang, Ritva Laury, Kumiko Ichihashi-Nakayama, Elise Kärkkäinen, Toshi Nakayama, Yoshi Ono, and Hongyin Tao for many inspiring discussions. Elise and Ritva and their families have become members of my extended family, with whom I can share things closest to my heart, be it linguistics or other fields of life.

Ritva Laury and Maria Vilkuna have both carefully read the manuscript for this book and provided many insightful comments. Tuomas Huumo and Lea Laitinen have read parts of this work and provided inspiring comments and important references. With Anneli Kauppinen, Tiina Onikki-Rantajääskö and Sara Routarinne I have been able to share my ideas even before they have received a decent shape. I am grateful to all of them for their interest in my work and their willingness to share their own work with me.

To my husband Mikko, I am grateful for sharing the adventure of life with me. With the help of his curious eyes I have been able to see things that I would have taken no notice of. I have greatly enjoyed watching our two sons Otso and Ilmo explore the world, and I am grateful to all three of them for giving my life a meaning that is far deeper than that of any academic accomplishment. My mother Anna-Maija Alanko has shown me how a woman can live a harmonious life both as a professional and as a mother. I am grateful to her as well as to my sister Kirsti Alanko for their unfailing support. From my father Aulis Alanko I learnt something that I feel has been important in my development as a linguist: in our long conversations about literature and music, he helped me to find my own voice so that I could learn to listen to what other people say. I dedicate this work to his memory.

CHAPTER 1

Introduction

What is a clause? How are clauses realized in conversational discourse? How do speakers attend to clauses as syntactic units when they are engaged in conversation? What other kinds of recurrent syntactic units are there besides clauses? With these research questions in mind, I have looked at discourse data from Finnish conversation in order to explore how syntax is being structured in discourse. I have studied syntactic structures in their local contexts to find out about the more global patterns and constraints on the use of these structures.

This study takes as its starting point the assumption that grammar is not a discrete set of rules that exist *a priori*, detached from the contexts of use, but rather, it emerges in discourse (Hopper 1987, 1988). Instead of being mentally and logically presupposed by discourse, grammar is evolving and being negotiated in discourse. This approach, known as “Emergent Grammar,” has been described by Hopper (1988) in the following way:

“[T]he “Emergence of Grammar” (—) attitude (—) has come to view grammar as the name for a vaguely defined set of sedimented (i.e., grammaticized) ‘recurrent partials’ whose status is constantly being renegotiated in speech and which cannot be distinguished in principle from strategies for building discourses.” (Hopper 1988: 118.)

My aim is to study the grammaticized ‘recurrent partials’ in their natural environments of use. I will analyze the local contexts where the grammaticized structures occur in order to show how they are being used and what kind of functions they serve in discourse. This way I hope to shed light on what these structures are needed for in discourse and, ultimately, why they have been grammaticized. I will also look at how the participants orient to them (e.g., whether they treat them in their response as complete or as projecting more) and thus, how they negotiate the status of the grammaticized items in the local contexts of use (see especially Chapter 4).

The term grammaticization as used in this work is thus not just referring to the historical process whereby a lexical item becomes grammatical or a less grammatical item becomes more grammatical (Heine et al. 1991: 3), but also, it refers to the more fundamental ontology of grammar, i.e., to the ways in which

grammar can be said to exist: grammar is “a vaguely defined set” of grammaticized recurrent patterns of discourse (see esp. Hopper 1987, 1988; cf. also Chafe 1994; Du Bois 1985, 1987a; Givón 1979; Ono and Thompson 1995). Thus, in order to study grammar we need to look at discourse and search for grammaticized recurrent patterns. (See Section 1.1 for further discussion.)

This study is not concerned with the history of the grammaticized items or constructions *per se*, but rather, attempts to describe their grammar as it is being realized in present-day Finnish discourse. However, the assumption that grammar results from the grammaticization of recurrent discourse patterns lays certain constraints on the synchronic description: it should not be in obvious conflict with what is known about the history of the linguistic element or construction in question.

The present study is data-driven in the sense that it tries to answer research questions that arise from the syntactic patterns found in the data. (For a discussion of the data and its validity for wider generalizations, see Section 1.2) The main focus will be on the analysis of Finnish conversations, but the results will be compared to cross-linguistic literature on syntax and discourse. Finnish is a language which has an elaborate case marking system to express grammatical relations, and I hope to shed light on the effects that this has on the discourse patterns. Furthermore, I will show that although Finnish is often described as a language with “free word order”, conversational discourse exhibits a strong tendency for a fixed order of subject and verb.

Looking at syntactic structuring from a conversation analytic perspective, Goodwin (1979) has proposed the following view on sentences, roughly corresponding to our notion of a clause (for further discussion, see Chapter 2):

Sentences emerge with conversation. However, *in traditional linguistics* it has been assumed that the analysis of sentences can be performed upon examples isolated from such an interactive process. In opposition to such a view it will be argued here that *sentences in natural conversation emerge as the products of a process of interaction between speaker and hearer* and that they mutually construct the turn at talk.” (Goodwin 1979:97–98; emphasis added by M. L. H.)

The goal of this study is to study clauses as they emerge “as the products of a process of interaction between speaker and hearer”. I will also show that there are other kinds of syntactic units in addition to clauses, such as free noun phrases (henceforth: free NPs) that can be used to construct turns. Since Goodwin’s (1979) paper, the field of discourse studies has seen a remarkable growth, and there is an emerging, alternative tradition called interactional

linguistics that looks at grammar in interaction, combining the findings of conversation analysis (see, e.g., Goodwin and Goodwin 1987, Jefferson 1990, Lerner 1991, Schegloff 1987, Sorjonen 1997) and discourse functional linguistics (see, e.g., Bybee and Scheibman 1997, Chafe 1994, Du Bois 1985, 1987a, 1987b, Duranti 1994, Ford 1993, Ford, Fox, and Thompson 1997, Fox 1987, Geluykens 1992, Ono and Thompson 1994, Ochs 1988, Tao 1996, and papers in Davis 1995, in Couper-Kuhlen and Selting 1996a, and in Ochs, Schegloff and Thompson 1996; on interactional linguistics, see Couper-Kuhlen and Selting in press).

The growing research interest in the interplay between language structure and language use as well as grammaticization phenomena has led to a critical scrutiny of basic grammatical units, such as the clause (see, e.g., Cumming 1984, Ono and Thompson 1995, Tao 1996). I take the clause to be an abstract grammatical unit, a type which is based on discourse tokens. It encompasses a predicate and its core arguments which may be accompanied by additional oblique arguments; however, I assume that the relationship between the predicate and its core arguments is grammaticized to a greater extent than that between the verb and the oblique arguments. In addition to the clause, speakers can use free NPs as a grammatical resource (see Chapter 3). Free NPs (also called “unattached NPs” by Ono and Thompson 1994) are characterized by the fact that they do not stand in a syntactic relationship to a verb; however, they do have internal syntactic structure, as do any other NPs (for example, a free NP may consist of a modifier and a head).

My central questions are the following: What are clauses and NPs as discourse constructs? What do speakers use them for? Chafe (1994:234) suggests that events and states are likely to be verbalized as clauses. In contrast, new participants are introduced into the discourse and tracked throughout the discourse by NPs (see Hopper & Thompson 1984 on the prototypical functions of nouns). To interpret these suggestions in terms of grammaticization processes, we could say that the discourse pressure felt by speakers to be able to describe events and states has given rise to constructions like clauses. The main participants of events and states, i.e., the ones that are being tracked in the discourse, are most typically given grammatical expression in a core role, while the circumstances and the minor characteristics of events and states are most typically coded as obliques. (See Chapter 2 for further discussion.)

Thus, this study takes as its starting point the view that grammar evolves in discourse through a process whereby the most basic and most common patterns of spoken discourse become crystallized, frozen forms that together

comprise grammar. In this view, grammatical relations emerge from discourse: recurrent patterns in discourse become grammaticized, and thus, part of the grammar. I will further assume that certain parts of grammar are relatively stable, while others may be more flexible.

1.1 Grammatical structures as cognitive schemas

In this section, I would like to address the question of the status of the grammatical constructions under discussion: what kind of constructs are they?

According to Langacker (1987:57), grammar is “a constantly evolving set of cognitive routines that are shaped, maintained, and modified by language use”. In his model, these cognitive routines are discussed in terms of schemas. In what follows, I will discuss schemas with respect to three characteristics that I consider crucial to my work:

- i. schemas as abstract prototypes,
- ii. schemas as templates for analogous expressions,
- iii. schemas as expressions of “socially shared” cognition (Schegloff 1991).

Ono and Thompson (1995) apply Langacker’s theory of constructional schemas in their discussion of conversational syntax. Ono and Thompson claim (1995: -217), “The evidence shows that syntactic constructions are attended to by conversationalists as abstract PROTOTYPES, and that speakers are highly tolerant of utterances which only roughly approximate any ‘constructional schema.’” They point out that the realization of schemas in conversation is heavily constrained by factors that are not grammatical *per se*, such as the position of the utterance in the sequential organization of the conversation and various cognitive constraints on memory and the like (see Schegloff 1979:281–282, Chafe 1980, 1987; for further discussion, see Chapter 5). Despite these factors, speakers are able to recognize utterances as instantiations of particular schemas, because they do not expect them to exactly correlate to the schemas.

Consider the following example which comes from a telephone conversation between mother and daughter (for a more specific description of the source of data see Section 1.2 below). The mother has been reading a catalog with various items for Christmas decoration, and has asked the daughter to buy something for her. At this point, she adds a further item for the daughter to buy.

(1) Mother and Daughter

- 1 Mother: *sit sä voisit ottaa, (H) kaks enkelii mul.*
 then you could take two angels I-ALL
 then you could take, two angels for me.
- 2 Daughter: *mitä enkelii.*
 What angels.
- 3 Mother: ... (1.2) *siel on kuule enkeli lumivalko, ... (0.3)*
 there is listen angel snow-white
ja pitsienkeli.
 and lace angel
 ... listen, there is Angel Snow White.. and Lace Angel.
- 4 *ja ne on molemmat, ... (0.6) kuus kahdenksankymment.*
 and they are both six eighty
 And they are both six eighty.
- 5 Daughter: *oukei.*
 Okay.
- 6 *[kumpaa] sä haluut.* ←
 which.one you want
 Which one do you want?
- 7 Mother: *[ja mää,]* ←
 and I
- 8 *e ... (0.4) no mielummi enkeli lumivalko,* ←
 er oh preferably angel Snow White
 er... well, Angel Snow White preferably.
- 9 *mut ei sil oo mitää välii kato ku,*
 but NEG it does nothing matter see because
 but it doesn't matter at all you see because
- 10 *mä annan ne, ... (0.4) ton, Ahosen Kaijal ja*
 I give them this FemName-ALL and
- 11 *Vatasen Liisal.*
 Fem Name-ALL
 I'll give them to ... this Kaija Ahonen and Liisa Vatanen.

In the analysis of the example, I will focus on lines 6–8, where there is some problem in speaker transition as can be seen in the speaker overlap. The overlap on lines 6–7 comes at a point where the mother has come to a possible completion of her request (line 4). The daughter gives a response that recognizes the

request (line 5), and goes on to make a further question about the request. Overlapping with the daughter's question, the mother starts a new utterance *ja mä* 'and I' (line 7) but cuts it off and answers the question (line 8). The utterance in line 7 is a possible beginning of a clause (connective + subject pronoun), but it is dropped in order to complete the question-answer sequence initiated by the daughter. Example (1) thus illustrates how the realization of schemas is constrained by the sequential organization of the conversation.

According to Langacker (1987:68, 1991:46), constructional schemas serve as templates for analogous expressions. The more frequent a pattern is, the more entrenched it is (Langacker 1987:59, Du Bois 1985, 1987a, 1987b). In terms of grammaticization processes, the more a given pattern is employed, the more conventionalized it becomes (see Heine et al. 1991 on principles of grammaticization).

I will now illustrate this with two examples from Finnish. The utterance in the first example realizes a transitive clausal schema with a highly grammaticized subject role, which can be defined in fairly abstract terms. In contrast, the second example illustrates a schema for an existential construction with lower level constraints on the realization of the different elements in the schema. (For more discussion, see Chapter 3.)

Consider the following example:

(2) Brothers

Pekka: *mä tempasi-n sielt semmose-n,*
 1SG caught-1SG from.there a.kind.of-ACC
 I caught a,
... yli puoltoist kilose-n raudu-n.
 over one.and.a.half kilo-ACC trout-ACC
 ... over one and a half kilos trout from there.

In example (2) the verb *tempasin* 'caught' has the 1st person pronoun *mä* as its subject. The subject pronoun is in the nominative and it triggers agreement in the verb (hence the 1st person marker *-n* on the verb). The accusative NP *semmosen yli puoltoist kilosen raudun* 'an over one and a half kilos trout' functions as the object. Thus, the clause in example (2) realizes an abstract schema for transitive constructions that could be described (in broad terms) in the following way:

NP V NP

There are certain constraints on the different elements in the schema. These constraints can be stated in abstract grammatical terms. *Inter alia*, the subject NP has to be in the nominative case and the verb shows agreement with the subject in number (singular : plural) and person (1st, 2nd or 3rd person). The object NP shows object case marking. The ordering of the constituents in the construction is constrained by discourse factors, although there is a clear preference for a subject–verb–object ordering (for more discussion, see Chapter 2).

This is in contrast with the schema for existential constructions. Consider example 3.

(3) Women

Linnea: *siel o-n syv-i-i koht-i-i*
 there be-3SG deep-PL-PTV spot-PL-PTV
 There are deep spots.

In example (3), there is a locative phrase *siel* ‘there’ in the beginning of the clause, followed by the verb *on* ‘be’ which is in the 3rd person singular form. After the verb, there is the partitive NP *syvii kohtii* ‘deep spots’, which is, according to the received view in Finnish linguistics, an “existential subject” (see e.g. Hakulinen and Karlsson 1979, Vilkuna 1996). However, it lacks the syntactic characteristics of subjects, namely it is not in the nominative case and does not trigger agreement (note that the E-NP is in the plural, but the verb is in the singular). The clause is a realization of the schema for existential constructions, which could be stated (roughly) as follows:

$NP_{\text{LOC}} + olla \text{ ‘be’} + NP_{\text{NOM/PTV}}$

In this schema, the first NP appears in a locative case, whereas the latter NP is either in the nominative or the partitive. The verb is usually *olla* ‘to be’ or sometimes some other verb with an existential meaning (such as *esiintyy* ‘occur’). The word order is quite fixed with the locative NP as the first element, the verb and the nominative/partitive NP, the so-called E-NP. In sum, there are lexical constraints on this construction. Furthermore, the ordering of the elements in the construction is specific to this construction type.

We could say that the subject role has been grammaticized in Finnish as a syntactic role which can be defined in abstract terms, without, say, lexical information about verb types or construction types, whereas the grammaticization of the E-NP has not proceeded that far, but is dependent on the grammaticization of the existential construction type as a whole. Bybee et al. (1994: 11) show that lexical or grammatical items undergo grammaticization change in

specific constructions. Thus, whole constructions, and not just simply the lexical meaning of the stem, function as sources of grammatical meanings.

If we want to characterize these grammaticization phenomena in terms of constructional schemas, we can see that they are entirely different: the schema for the subject is abstract and general, whereas the schematic information concerning E-NPs contains information which is specific to the construction type, i.e. existential constructions, and which sets constraints on the order of elements, case marking of the nominals and choice of verb.

The instantiation of a particular schema may be partial, and it may differ from the schema with respect to a certain parameter. This may function as the motivating force for syntactic change: the instantiations, if they are frequent enough, may serve as a basis for new schemas (see Nichols and Timberlake 1991). Thus, the schemas that speakers have do not form a finite set, but are best understood as a dynamic inventory.

Ono and Thompson suggest (1995: 13) that syntactic knowledge is a form of “socially shared cognition” (see Schegloff 1991); “what we ‘know’ cannot be separated from what we ‘do’”. It seems to me that an assumption like this is necessary in order to account for the cognitive structures that lie behind the use of language (especially if we reject the view that linguistic knowledge is innate, as is done in the present work); otherwise it would be hard to explain how speakers can be said to speak the same language.

In his article on socially shared cognition, Schegloff (1991: 168) claims that the Western cultural tradition is heavily biased in its emphasis on the single individual. He blames the cognitive sciences for studying cognition “in the splendid isolation of the individual mind or brain” and for reserving “the social aspect for later supplementary consideration”. Schegloff suggests that this is misconceived and claims that “our understanding of the world and of one another is posed as a problem, and resolved as an achievement, in an inescapably social and interaction context” (1991: 168).

With the methodology applied in my work, the only way we can approach cognitive structures like schemas is through a thorough analysis of the data: the methodology allows us to make observations on the patterns that emerge from discourse data (see esp. Du Bois 1987b). These patterns form the basis for the assumption of more abstract categories on an emic level (such as NP schemas, clause schemas).

In conclusion, I assume, following Langacker (1987: 57), that grammar is “a constantly evolving set of cognitive routines” which is shaped and maintained through language use. These cognitive routines are assumed to represent

“socially shared cognition” (Ono and Thompson 1995). The routines can be described in terms of schemas, which serve as templates for analogous expressions.

1.2 Data

This study is based on naturally occurring conversations between speakers of Finnish. Following many researchers both in discourse linguistics and conversation analysis, I take conversation to be the natural habitat of language, representing the most fundamental condition of language use. Furthermore, it is the element in which grammaticization patterns can be most readily observed. In this section, I will introduce my data.

The data for this study come from 6 audio-taped conversations between speakers of Finnish, together amounting to approx. 6 hours of audiotapes. From the conversations, I have chosen smaller segments for closer analysis, approx. 5–10 minutes each, together comprising some 40 minutes of conversation. This is the core database that contains 1981 intonation units (see Section 1.4 for a discussion of intonation units). The segments chosen for closer analysis include both narrative conversational sequences and more rapid exchanges. In what follows, I will present each conversation with information on speakers and their backgrounds. There will also be a brief characterization of the conversation itself and the topics being discussed.

Birthday is a conversation between 7 young people, all students at the University of Helsinki. However, only one of them comes from Helsinki; the others come from different parts of the country. They are having dinner together and chatting. The conversation contains several series of stories, concerning inter alia travel experiences, people’s first visits to their boy- or girlfriend’s parents’, and cats and animal in general. Between the stories there is talk related to eating and drinking and gossip. The tape was made in 1989 and belongs to the corpus of spoken interaction at the department of Finnish, University of Helsinki.

Women is a conversation between four middle-aged women who work in the same pharmacy. The daughter of one of the women is also present, and while making the recording, she is occasionally commenting on the older women’s talk. The topics range from summer cottages to work, from parties to travel experiences. The speakers come from South-Western Finland. The recording was made in the 1970’s for the corpus of spoken Finnish at the University of Turku.

Alko is a conversation between four men who are all work mates. The data were originally collected in the 1950's for a sociological study where the participants were given different kinds of tasks (Bruun 1958). However, the segment that I have chosen for analysis was recorded after the participants had finished a task and were free to discuss anything they wished. They discuss drinking and smoking habits, the cost of living, housing in Helsinki, and kids. The recording takes place in a banquet room that is adjacent to a restaurant, and the participants are offered free food and drink. All speakers live in the Helsinki area, but they come from different parts of the country, and there are dialectal differences that can be heard in their talk. The tape belongs to the corpus of spoken interaction at the department of Finnish, University of Helsinki.

Family is a conversation between an adult daughter and her parents, who live in Turku, South-Western Finland. The parents have just come home from the theater and they are discussing the play and gossiping about the people in the local theater. The recording belongs to the corpus of spoken Finnish at the University of Turku.

Mother and daughter is a telephone conversation between mother and daughter. The mother is asking the daughter to run some errands for her, but she also tells about her day. The daughter is gossiping about a mutual friend and her family. The mother and daughter both live in Helsinki, but the mother comes from South-Western Finland. The recording was made in 1989.

Brothers is a telephone conversation between two brothers. The other one has just been to Lapland in Northern Finland on a fishing trip and is calling his brother to tell about the trip. The tape was made in 1990 and is part of the corpus of spoken interaction at the department of Finnish, University of Helsinki.

All conversations in the data can be characterized as being informal everyday interactions where the participants know each other well. The data do not cover all the different dialectal areas in Finland, and they are also slightly skewed towards younger generation of speakers. However, the existing literature in Finnish linguistics does not point to potential dialectal differences in the phenomena studied here, namely clausal relations and free NPs. Those dialectal features that can be heard for example in the *Alko* conversation are morpho-phonological in nature.

I have analyzed the core database word by word, phrase by phrase in fine detail (see the following section for discussion of the coding). In addition to this database, I have collected a database of free NPs by listening to tapes from

approx. 6 hours of multiparty conversations and transcribing each instance of a free NP with relevant context. This was done to secure the reliability of the analysis of free NPs, since some types of free NPs (such as constructions with a theme and orientation) are relatively infrequent in the data.

The focus of this study is on examining syntactic structuring in Finnish. At the same time, I have tried to watch for any possible cross-linguistic implications by comparing my analysis with studies from different languages. As I will show, there indeed is evidence for interesting differences and similarities across languages in the ways in which syntax is put to serve various interactional goals.

1.3 Coding and analysis of the data

As I mentioned in the previous section, I have done detailed coding of the core database. In this section, I will discuss the principles that I have followed in the coding.

1.3.1 Noun phrase type

Each NP was coded as either pronominal or lexical (also called full NP), and among pronouns, personal pronouns were further coded for person (1st, 2nd or 3rd). It has been shown in cross-linguistic studies that personal pronouns often exhibit coding patterns that are distinct from those of other pronouns and nouns. This is true of Finnish also, and in Chapter 2, I will discuss the implications of these differences to the description of syntactic coding patterns.

Number of the NP (singular vs. plural) was also coded, as well as case of the head of the NP. As will be discussed in more detail below, modifiers usually precede their heads in Finnish and agree with the head in case and number. Consider the following example.

- (4) Brothers
me oltiin yhe-s semmose-s joki-uoma-ssa
 we were one-INE such-INE river-bed-INE
 We were camping in one such river bed.

In example (4), the whole NP indicating location is inflected in the inessive case — the modifiers *yhes semmoses* ‘one such’, as well as the head *jokiuomassa* ‘river bed’.

1.3.2 Semantic class

Each NP was coded for the semantic class of the referent. The distinction between human vs. non-human referents turned out to be significant here, especially in relation to the observed word order patterns (see especially Chapter 2), and the characteristics of the subject role in different clause types (see Chapter 3).

1.3.3 Activation cost

According to Chafe (1987, 1994), a particular concept may be in one of three different *activation states*, active (given), semiactive (accessible), or inactive (new) at a particular time in conversation, and respectively, in focal, peripheral, or inactive consciousness. To illustrate these three states, Chafe uses a metaphor taken from trade and talks about activation cost. The idea behind activation cost is that a given item is least costly cognitively because it has already been active, whereas a new item is more costly because it requires more mental energy to convert it from the inactive state to the active state (Chafe 1994: 71–75).

Many studies have associated new information with prosodic accent so that new mentions have been found to be accented (i.e. they carry primary or secondary accent (see e.g. Chafe 1976, 1994, Prince 1981, Terken and Hirschberg 1994; for further discussion, see Chapter 4). According to Chafe, given mentions are typically verbalized with a weakly accented pronoun (at least in English; Chafe 1994: 81).

Discourse studies on different languages have indicated that there are interesting correlations between noun phrase type and the activation status of the referent (see e.g., Lambrecht 1985, Du Bois 1985, 1987b, Chafe 1994).

In my data, I have limited the coding of activation cost to indicating whether the referent was new to the discourse or not. This was done because there is reason to believe that it is the marked category, i.e., new mentions, that shows the most interesting patterns (cf. Larsen 1981, Du Bois 1987b, Herring 1989, Bentivoglio 1992, Ashby 1995, and Ashby and Bentivoglio 1993). A NP was coded as new if the referent had not been mentioned before and it was not immediately present in the context.

1.3.4 Discourse referentiality

One central function of NPs is to allow speakers to speak about an entity as having “continuity of identity” (Du Bois 1980: 208–209), or, in other words, to track that entity through discourse (Du Bois and Thompson 1990). According to Du Bois and Thompson (1990: 24), *tracking NPs* either introduce a new participant in the discourse or serve to track one.

Durie (1994: 504) discusses a similar notion under the name *trackable* (vs. *non-trackable*). According to him, trackable referents are “of the kind that could be mentioned again”. Furthermore, trackables require “the hearer to open a file for them”.

In my coding procedure, I have not considered whether a referent *could* be mentioned or not, but instead, I just considered whether it really was mentioned again. Thus, in analyzing my data, an NP was coded as tracking if it was mentioned more than once in the discourse.

1.3.5 Grammatical roles

The data have been coded for grammatical role: subject, object, predicate nominal, oblique, free NP, and E-NP (see Section 1.1 above, and Section 3.1.1). In this section, I will first discuss the category of subject briefly (for more discussion, see Chapter 3). Then there will be a discussion of the analysis of the object in passive clauses in Finnish.

In current typological literature it has become customary to refer to NPs in the clause core with the labels A, S (or sometimes P) and O (see e.g. Dixon 1972, 1979, Silverstein 1976, Comrie 1981, Blake and Mallinson 1981, Blake 1994; for a critical discussion, see Chafe and Mithun 1995). What the labels are taken to mean, seems to vary from researcher to researcher. Dixon (1979: 108) takes them to represent “syntactic-semantic functions”, and describes S as the only obligatory NP in an intransitive clause, A as “the NP in a transitive clause which can be agent”, and O as “the other obligatory NP in a transitive clause”. These definitions contain a blend of syntactic (e.g., obligatory NP in an intransitive clause) and semantic characterizations (agent). Elsewhere in the same article Dixon (1979: 59) gives a straight-forward syntactic definition, according to which A is the transitive subject, S the intransitive subject and O the transitive object (see also Du Bois 1987a and 1987b).

Originally, Dixon (1972) introduced the labels A, S, and O as a way of explaining the functioning of an ergative marking system to those not acquainted

with such a system, and since then, the labels have been put to use in several studies exploring ergativity, especially coding splits attested ergative languages (see e.g. Silverstein 1976, Du Bois 1985). While admitting that A, S, and O may be useful when describing ergative languages in comparison with nominative-accusative languages, I have chosen not to use these labels when describing the patternings I have found in my Finnish data. Instead, I will speak about subjects and objects. As I will show in Chapters 2 and 3, the subject is a useful category in Finnish grammar and discourse.

In his studies on Sakapulteko spoken narratives, Du Bois (1985, 1987a, 1987b) has found that transitive subjects (Du Bois' A's) carry out discourse work that is very different from that of intransitive subjects (S's) and objects (O's; see §2.4 for further discussion). Namely, transitive subjects are rarely expressed with full NPs, whereas intransitive subjects and objects are more likely to be full NPs. Furthermore, transitive subjects seldom bring new referents into the discourse, whereas intransitive subjects and objects are much more likely to contain new information. In other words, transitive subjects show a discourse profile that is quite different from that of intransitive subjects and objects. Du Bois concluded that Sakapulteko shows ergative alignment not only in grammar but also in discourse patterning. (Du Bois 1985, 1987a.) Du Bois' studies have generated a wealth of research into the relationship between argument structure and information flow in various languages (see literature sighted in Du Bois 1987a, Durie 1988, Ashby and Bentivoglio 1993, Kumpf 1992, Kärkkäinen 1996, Du Bois et al. forthcoming). With this line of research in mind, I coded transitive and intransitive subjects separately (St and Si, respectively) so that I could see whether they show different discourse profiles in Finnish. As I will show in Chapter 3, they indeed do, but not quite in the way Du Bois originally suggested. I have a further coding Sc for subjects of copular clauses.

In Finnish, the object is usually distinguished from the subject with case marking: while the subject is in the nominative, the object is marked with the accusative or the partitive. However, if there is no subject in the clause, and thus, no need to distinguish the object from the subject, the object can stand in the nominative. This alternation will be discussed more closely in Chapter 2. Here I would like to draw attention to the fact that in Finnish the passive does not function to change the analysis of the object. This can most clearly be seen in clauses where a personal pronoun functions as the object. Consider the following examples. Example (6a) is an active clause with a personal pronoun object *meidät* 'we-ACC'. Example (6b) is fabricated.

- (6) a. Brothers
 (se) *vet-i* *meidä-t mukana sinne ylös.*
 3SG pull-PAST+3SG we-ACC with there up
 He pulled us up there with (his snow mobile).
- b. Fabricated example
meidä-t vedet-t-i-in *mukana sinne ylös.*
 we-ACC pull-PSS-PAST-PERS with there up
 We were pulled up there with (a snow mobile).

The object pronoun *meidät* ‘we-ACC’ is in the accusative form in example (6b) despite the fact that the verb is in the passive.

1.3.6 Summary

In conclusion, the data have been coded with respect to several morpho-syntactic features and information flow parameters in order to find out about patterns that emerge in the discourse data. However, the coding should be seen as a first step in the analysis which is to be followed by interpretation. In the discussion of the following chapters some frequency counts will be given that are based on the codings, but they are used only as guides in the search for a functional explanation for the observed patterns. By no means do they constitute explanations themselves.

1.4 Representing the data

I have transcribed the data myself. For some of the segments, I was able to check my transcription against somebody else’s (this was true of the Birthday party conversation, the Alko conversation and parts of the Women conversation). The transcriptions were initially made following the conventions of Du Bois et al. (1991, 1993, see appendix for a list of the conventions), but when I checked the transcription of the segments that I had chosen for the examples, I made changes in the transcripts to accommodate the specific needs of the context of discussion where a given example appeared. I left out some of the details which were not relevant to the point that was being discussed, and highlighted some other features that had turned out to be important to the issues under discussion.

In the Du Bois et al. transcription system, special attention is given to intonation, and the data is transcribed in terms of intonation units. This is of course not something peculiar to this system of transcription, as many studies assume that connected speech can be divided into smaller units defined in terms of intonation, or more specifically, in terms of intonational contour. These units have been called by various names, including intonation-groups (Cruttenden 1986), tone-groups (e.g. Halliday 1967), intonational phrases (Pierrehumbert 1980, Couper-Kuhlen 1996) and intonation units (Chafe 1987, 1994; Du Bois et al. 1993), all roughly equivalent to each other, at least in principle. There are considerable differences between respective approaches: for example, Pierrehumbert's theory of intonational phrases is based on acoustics, whereas the others are auditory approaches. Also, Cruttenden's intonation-groups and Pierrehumbert's intonational phrases function in the domain of a single sentence, whereas Halliday's approach as well as that of Chafe and his associates are oriented towards discourse. Couper-Kuhlen (1996) and Auer (1996) combine acoustic and auditory analysis in their discussions of intonation in a conversation analytic framework.

In my work, I have made use of the term intonation unit, defining it as a "stretch of speech uttered under a single coherent intonation contour" (Du Bois et al. 1993). To help the identification of intonation units, there are additional cues: at intonation unit boundaries, there is often pitch reset in the beginning of the unit, acceleration in tempo on the initial unstressed syllables (anacrusis), pausing and prosodic lengthening of the final syllables (see Du Bois et al. 1993, Cruttenden 1986). Thus, intonation units have been defined as perceptual units. However, I used acoustic analysis to help the analysis in unclear cases.

In the transcripts, I marked transitional continuity making three distinctions: final intonation contour, characterized by a final falling contour, was marked with a period, continuing intonation with a slight fall in the contour was marked with a comma, and question intonation, characterized by a rising contour, was indicated with a question mark. In the examples, I have made note of voice quality and primary accent when that has been relevant to the point that I have wanted to make with the example.

In order to help the reader to follow the interaction in places with overlap and other interactionally sensitive points, I have chosen to represent the data so that intonation units do not always form a line all by themselves (as is conventionally done in the Du Bois et al. 1993 system), but rather follow each other. The following example illustrates this.

- (7) Alko
- 1 Aarre: *mu-ll on suuremma-t ^meno-t-ki tullee-vat*
 1SG-ADE is larger-PL expense-PL-CLT come-3PL
 I have more expenses / the expenses become more
- 2 [*helevetti. et,*]
 damn.it that
 damn it. So,
- 3 Pekka: [*joo, näin on.*]
 yeah so is
 Yeah, you are right.

In the example, lines 2 and 3 both contain two intonation units. In line 2, there is an intonation unit boundary between *helevetti* and *et*, as can be seen from the period that marks the transitional continuity of the first intonation unit. Note that the unit ending in *helevetti* starts in the beginning of line 1. This can be read from the transcript from the fact that there is no mark for transitional continuity before the period after *helevetti*. The first intonation unit in line 3 ends in a continuing intonation, and is followed by an approving comment *näin on*, ending in a final intonation contour. By representing the two intonation units on the same line, the overlap between Pekka and Aarre is easier to grasp from the transcript.

The transcription is only a poor image of the actual interaction, and reflects the transcriber's theoretical preferences and interests. In my view, however, we can make the transcription a useful tool that can be used to search for interesting cases and parallels, but it is important to bear in mind that transcripts are never final.

1.5 Organization

The discussion in the rest of this book is organized as follows: Chapter 2 discusses clauses and clausal syntactic relations. The main research question in that chapter is what clauses are as discourse constructs. In Chapter 3 we will discuss the subject role from the viewpoint of the grammaticization of syntactic functions. Chapter 4 provides a look at free NPs, their syntax and discourse functions. Finally, Chapter 5 discusses the functions of intonation in syntactic structuring. In Chapter 6 I will draw some conclusions.

CHAPTER 2

Dynamics of the clause

Compared to normative written language, the grammar of spoken language may seem chaotic and difficult to describe. Even the basic terms of syntactic description may seem questionable. Consider this call for participation (“Workshop on Spoken Dialogue at ECAI (European Conference on Artificial Intelligence”), posted on the Linguist list 2/15/96), in which Susann LuperFoy laid out the following problem for the development of dialogue systems for speech:

“While for dialogue systems that cope with written/typed language a sentence can serve as basic unit for dialogue modeling, this approach cannot be applied for the treatment of spontaneous speech or even read speech, where phrasing of sentences may differ from speaker to speaker. Spoken input is often incomplete, incorrect and contains interruptions and repairs; full sentences occur only very occasionally. Therefore, new basic units for the development of dialogue models have to be proposed in order to also capture fragmentary input.”

I find this statement extremely interesting for many reasons. First, it assumes that the concept of sentence is unproblematic as a basic unit in written language, but problematic for spoken language. Thus, spoken language differs from the written language in important respects. Secondly, it is the spoken language that is deviant: it is incomplete, incorrect, and contains interruptions and repairs. Therefore, new basic units are needed to “capture fragmentary input”. In my view, these problems are, for the most part, created by the viewpoint: if we take standard written language as a starting point, many of the features that characterize spoken language may seem problematic just because they do not exist in written language. However, if we take spoken language as our vantage point, it is easier to develop tools to describe the characteristics of spoken discourse, and it is my belief that the very same tools will handle many features of written discourse also.

Furthermore, the received view of the characteristic features of the written language appears to be idealized in many respects. For example, it is not difficult to find examples in fiction where typed sentences do not match normative sentences (8) or where what is between full stops is not a sentence

but another kind of a unit, a free NP (9; in the example, The Maginot Line is a nickname to a woman).

- (8) Toni Morrison, *The Bluest Eye* p. 86

She liked, most of all, to arrange things. To line up things in rows — jars on shelves at canning, peach pits on the step, sticks, stones, leaves — and the members of her family let these arrangements be.

Ex. 8 starts out with a clause with an infinitival complementizer *to arrange things*. After the full stop, there is another infinitival construction *to line up things in rows*, which is in a paratactic relationship with the infinitival construction in the first clause, although there is no formal marking for this relationship. Between the hyphens, there is an insertion, a list of *things* the person likes to line up. After the insertion, the conjunction *and* forms a coordinating relationship between the last clause and something else. It is not entirely clear what that something else is. One possible interpretation is that the first clause and the infinitival construction, together with the insertion, form a loose unit that the conjunction *and* coordinates with the last clause *the members of her family let these arrangements be*. However, the author has left this interpretation up to the reader.

- (9) Toni Morrison, *The Bluest Eye* p. 81

The Maginot Line put a fat hand on one of the folds of her stomach and laughed. At first just a deep humming with her mouth closed, then a larger, warmer sound. Laughter at once beautiful and frightening. She let her head tilt sideways, (...).

In example 9, there are three free NP constructions, the first two ones *a deep humming* and *a larger, warmer sound* describing the kind of sound the person made when laughing. The third one *laughter at once beautiful and frightening* further characterizes the sound and the emotions it arouses (for poetic effect, the modifiers *beautiful* and *frightening* appear after the head).

Thus, both (8) and (9) contain examples of cases where there is a clear conflict between typed sentence and normative sentence. It is important to note that cases like the ones above appear not only in literary fiction, but also in other genres as well. For example in scientific texts, titles and headings and captions to tables and figures often take the form of free NPs.

In recent years, there has been a growing research interest into the orderliness of spoken discourse, both within conversation analysis (see e.g. Sacks et al. 1974, Sacks 1992a and 1992b, C. Goodwin 1981, Jefferson 1990, Lerner 1991,

Schegloff 1979) and in discourse linguistics (see e.g. Chafe 1987, 1994, Ford 1993, Fox 1987, Lambrecht 1987, Tao 1996, papers in Davis 1996, Ochs, Schegloff and Thompson 1996, and Couper-Kuhlen and Selting 1996). *Inter alia*, Schegloff (1979) and Fox, Hayashi and Jasperson (1996) have shown that repairs are not chaotic and random, but show a clear patterning both in terms of what is repaired and how it is done. Instead of deeming spoken discourse “incomplete” and “incorrect”, like LuperFoy in the above statement, we should study it in its own right. The present study attempts to contribute to a functional theory of syntax by looking at how syntax is structured in spoken discourse.

This chapter will start off with a discussion of the structure of the clause as it has been conceived of in the functionalist literature. I will first discuss what I call grammar-to-discourse models, and then I will present another approach that proceeds in the opposite direction, namely from observed discourse patterning to grammatical structures. My aim is to present a definition of the category of clause relevant for the analysis of both spoken and written discourse. I will discuss different syntactic coding strategies, namely case marking, agreement and word order, and I will show that the clause core is where syntactic relations emerge most clearly. By the clause core, I mean the predicate and its core arguments, subjects and objects. The clause core may be accompanied by additional adjuncts, but they do not constitute the clause in the same way as the predicate and its core arguments do. We will also look at the management of information flow, and will see that the core arguments do very different kind of work from obliques in managing information and keeping track of participants in discourse.

2.1 What is a clause?

In the literature on discourse and grammar, the definition of the clause has not received much attention. Rather, it has been taken to have a shared, non-controversial meaning that needs no discussion. In this section, I will address the question of what clauses really are, and what they are needed for in grammar and in discourse. I will first review existing work on clauses in functional linguistics, and then continue to outline my own approach. The scope of the discussion is limited to functional literature because the research goals are similar to the ones adopted here, thus allowing for fair comparison.

2.1.1 Grammar-to-discourse models

The last two decades have seen a growing research interest in the interplay between language structure and language use. Observation of discourse patternings have shed new light on grammatical structures, and this has led to critical scrutiny of basic grammatical concepts, such as the clause. In this section, I would like to discuss the definition of the clause in what I call “grammar-to-discourse models”.

In his “function-based” introduction to English grammar, Givón (1994:22) offers a semantico-pragmatic definition: “Clauses, also called sentences, code propositions. A proposition combines concepts — i.e. words — into information. Information is about relations, qualities, states or events in which entities partake.” (See also Givón 1984:85.) For Givón, clauses are syntactic expressions of propositions, which in turn are semantic in nature. In a proposition, concepts are combined to produce information. What Givón seems to mean by information is the pragmatic interpretation of propositional content. Givón’s definition takes language structure (clauses) as a starting point, and then proceeds to semantic content (propositions) and pragmatic function (information).

In their functional grammar, Foley and Van Valin (1984) argue for a more structural approach, according to which clauses are seen as the building blocks of discourse. They state (1984:27): “Regardless of the type of discourse under consideration, the clauses which constitute the discourse are constructed around predications consisting of a predicate and its argument(s).” In this view, the center of the clause is the nucleus with the predicate. The next layer out is the core with the main arguments, whereas the periphery contains optional arguments. In sum, according to Foley and Van Valin, discourses are made up of clauses which in turn consist of predicates and their arguments. For Foley and Van Valin, discourse is yet another level of organization on top of the clause.

As Tao (1996:16) notes, this predicate-centered view of the clause is problematic since it is not clear what the notion of predicate is assumed to encompass. I will illustrate these problems with example (10), which is the closing sequence of a set of stories which Mella has told about her parents who travel around and always get into trouble.

(10) Birthday

- 1 Mella: *ne on ^kiertä-ny vaikka kuinka*
 3PL be+3SG go around-PCP just how much
pal[₁jon₁] mutta,
 much but
 They (Mella's parents) have traveled around a lot but
- 2 (X): [₁⟨X ole-X⟩ —₁]
 be
 Be —
- 3 Mella: *ne on niin [₂ uuno-j-a kun ←*
 3PL be+3SG as dumb-PL-PTV as
ol-la vo-i₂].
 be-INF can-3SG
 they are as dumb as anybody can be.
- 4 Sanna: [₂ *tyttäre-n kunnioitta-va a₂*] *sen[₃ne₃] ←*
 daughter-GEN honor-PCP attitude
 A daughter's respectful attitude.
- 5 (X): [₃mm₃].
 Uh huh.
- 6 *^niin.*
 so
 Right.

Focus on lines 3 and 4 and compare them. In line 3 we have a predicate nominal clause. In Finnish, there is good reason to analyze predicate nominal constructions as clauses: they contain the copular verb (*on* 'be', line 3) which together with the predicate nominal forms the predicate (*on uunoja* 'are dumb' on line 3).¹ The copular verb carries verbal trappings such as person marking and tense. In addition to nominal clauses, there are predicating NPs that are not part of any clause, but rather, they stand as free syntactic units, so-called free NPs (or unattached NPs; see Helasvuo 1988, 1991; Herring 1989; Tao 1992; Ono & Thompson 1994, Durie 1994). In line 4, we have an example of a free NP.

1. 'Dumb' is only a near equivalent of *uunoja*. *Uunoja* is etymologically not an adjective, but a proper noun, an old-fashioned male first name, which is here used as a characterizing term. In Finnish, speakers of a certain social or age group may use some name (usually a male first name) as a characterizing term, usually with a pejorative meaning, but after a while the usage is dropped.

The free NP is predicating, but in a very different way from the predicate nominal clause on line 3. In line 3, the 3rd person plural pronoun *ne* functions as a subject, and the copular verb *on* agrees with it in person (but not in number, see Section 2.2). The predicate nominal *uunoja* ‘dumb’ serves to characterize the referent of the subject pronoun, namely the parents. The free NP on line 4 is also characterizing, but unlike the predicate nominal in line 3, it is not syntactically related to any NP or verb. It serves to characterize, but it is not clear what exactly is being characterized: it could be Mella’s assessment in line 3, or the whole set of stories she has just told. As I will show in Chapter 4, it is typical of characterizing free NPs that they are used in contexts where the scope of the characterization is not specified.

In sum, we see from example (10) that clauses are not the only syntactic units that need to be accounted for in the analysis of spoken discourse. We also see that the relationship between the verb and its core arguments is not the only source of syntactic units in spoken discourse, but the relationship between a modifier and its head as realized in a free NP may create a syntactic unit independent from any clausal structure.

Some scholars (e.g. Halliday 1985, Herring 1989) have proposed that free NPs should be described as verbless clauses (or “minor clauses”). However, I take this to be misleading since free NPs lack many of the characteristics that define clauses (see Chapter 4). Most notably, there is no verbal predicate, and hence, no argument structure in free NPs. In a way, some characterizing free NPs may seem to resemble predicate nominal clauses in languages where copular clauses lack a copular verb (e.g. in Russian) and the predicate nominal or adjective serves as the predicate. However, even in these languages, there is normally some way to express tense and aspect and other verbal markings if they differ from the unmarked (for example, in Russian, although there is no copula in the present tense, the verb ‘to be’ is being used in the past tense). In free NPs, there is no possibility of expressing these verbal codings.

The view that discourse is made up of clauses is usually taken for granted rather than argued for. In writing, clause boundaries are most often marked typographically (with dots, question marks, exclamation marks etc.). In speech, the cues which indicate that one syntactic construction has come to an end and another one is starting are much more subtle. The cues may be syntactic: for example, if the core argument positions of a certain verb token have been filled, that certainly constitutes a cue for the co-participants that the clause has come at least close to its end. If an NP is semantically congruent with a verb close to it, we are likely at least to try to interpret it as an argument

of that verb (see analysis of example 10 below). In addition to argument structure, intonation may function as a cue. As I will show in Chapter 4, a final intonation contour counts as a strong indication that a syntactic construction has come to an end (see also Schuetze-Coburn 1992, Ford and Thompson 1996). However, there are many more syntactic boundaries than there are final intonation contours. Moreover, clauses are not the only kind of syntactic units that need to be accounted for in the analysis of spoken discourse.

However fascinating the building block theory of discourse as being made up of clauses may be, it runs into difficulty when we analyze real spoken discourse. To take the challenge of analyzing spoken discourse seriously, we may be forced to set aside the structuralist ideal of dividing everything into neatly distinguishable, clear-cut grammatical units.

2.1.2 What are clauses needed for in discourse?

In this section, I would like to outline an alternative approach to what clauses are. I will take two observations as a starting point: First, it is often not clear where one clause ends and another one starts in a stretch of spoken discourse. Thus, clauses are not clear-cut units but rather, emergent syntactic structures (cf. Hopper 1987, 1988). Secondly, free NPs and particle chains also form syntactic units that have equal validity with clauses. I will first give an overall outline of my approach, and then discuss these two points in more detail.

In an alternative approach, one would start with the discourse needs speakers have: speakers need grammatical structures — *inter alia* clauses — to serve certain communicative and interactional goals. Language and linguistic structure is not a means for self-expression by individual speakers, but for interaction among participants. An important but yet quite neglected aspect of language use is its reflexivity: language is used to structure discourse and interactional activities.

I argue for a view which sees clauses as emergent syntactic constructions consisting of the verb and its arguments. In addition, there may be oblique arguments which accompany the verb but which are not required by the verb. These oblique arguments belong to the clause but they do not constitute it in the same way as the verb and its core arguments do (for similar views, see Du Bois 1987b, Tao 1996: 17). Thus, by clauses, I mean the predicate verb and its core arguments, with optional oblique arguments. I include copular clauses under this definition, but exclude free NPs.

I will illustrate what I mean by emergent syntactic constructions with the

following example from the birthday party conversation. Before this point in the conversation, Sanna has told the other participants that she doesn't have the cheese she was supposed to bring to the party, although she remembers buying it and having paid for it; somehow she lost the cheese without noticing it. For background it might be helpful to know that Sanna, Noora and Henna share an apartment (cf. Noora's line 9) and that the party is at their place.

(11) Birthday

- 1 Mella: ... *ei ne oo su-n taki-n tasku-ssa*,
 NEG+3SG 3PL be 2SG-GEN coat-GEN pocket-INE
 They are not in your coat pocket, are they? (i.e. the pieces of cheese that Sanna was supposed to bring)
- 2 Noora: ... *mä kato-i-n ne-kin [jo]*.
 1SG look-PAST-1SG 3PL-CLT already
 I've already checked them also.
- 3 Sanna: [e=i],
 NEG+3SG
 Oh no,
- 4 *ai sä kat-*
 oh 2SG look-
 you checked- —
- 5 (:): @@ [1 @ @1]
- 6 Leena: [1 <Pmm. P>
 Uhhuh.
- 7 ... <P ^saappa-an [2 varre-ssa. P] >]1 ←
 boot-GEN-stem-INE
 In the bootleg,
- 8 Noora: [2 ei si-llä2]
 NEG+3SG it-ADE
 It's not
- 9 [3 et mä yrittä-isi-n jotenkin ^holho-ta3]
 that 1SG try-COND-1SG somehow patronize-INF
 that I would try to patronize
- 10 Sanna: @[3 <HI tää mene-e liian paha-ks HI>3].
 this go-3SG too bad-TRA
 This gets just too bad/crazy.

- 11 Noora: [₄ *näi-tä mutta?*₄]
these-PTV but
these guys somehow but,
- 12 (Leena): [₄ *ohhoh ohhoh*₄]
Oh no!
- 13 Henna: *<P otta-kaa P> leipä-ä.*
<P take-2PL bread-PTV
Take some bread!
- 14 (X): *ahmm=*
- 15 Sanna: @ [₅(*H*)₅]
- 16 Leena: [₅ *^saap*₅]*paa-n-varre-ssa,* ←
boot-GEN-stem-INE
In the bootleg.
- 17 Sanna: *oh*[₆ *hoi*₆]*jaa.*
Oh dear.
- 18 (Leena): [₆ *mm.*₆]
Uhhuh.
- 19 (Leena): *mm.*
Mm.
- 20 Noora: *si-tä mä e-n katto-nu.*
3SG-PTV 1SG NEG-1SG look-PCP
That one I didn't check.
- 21 (X): (*THROAT*)
- 22 Leena: *nii siel* [₇ *voi ol-la*₇].
so there may+3SG be-INF
So (it) could be there.
- 23 Raita: [₇ *<P syylingi-ssä? P>*₇]
felt.sock-INE ←
In the felt sock? (Dialectal word)
- 24 Leena: [₈ *mikä*₈].
what
What?
- 25 Sanna: [₈ *ne ol-i*₈] *aika< iso-t pala-t >*
3PL be-PAST+3SG quite big-PL piece-PL
They were pretty big pieces.

In this sequence, the participants are engaged in trying to locate the cheese that Sanna has lost. The sequence starts out with a question (line 1) by one of the guests, Mella. The question is formulated as a negative statement *ei ne oo sun takin taskussa* ‘they are not in your coat pocket (are they?)’.² Noora gives an answer (line 2) that implies that the pockets are not the only place that she has already checked. This implication is indicated through the use of the clitic *-kin*, translated here as ‘also’. Throughout the passage, Noora presents herself as a person who looks after her roommate Sanna (cf. lines 2 and 20 and the ironic statement on lines 8, 9 and 11), an attitude which is not readily accepted by Sanna (cf. lines 3, 4, and 10). Leena suggests yet another place to be checked in line 7 *saappaanvarressa* ‘in the bootleg’. This suggestion is made with a low voice in the midst of a roar of laughter and it is partly overlapped by Noora (line 8). As Leena’s suggestion is not taken up by the co-participants, she repeats it in line 16 and manages to elicit responses from Sanna (line 17) and Noora (line 20). Raita adds one more suggestion, the humorous *syylingissä* ‘in the felt sock’ (line 23) which is no more realistic, but rather, an exaggeration of Leena’s suggestion. Sanna’s utterance *ne oli aika isot palat* ‘they were pretty big pieces’ (line 25) is a comment on all the suggestions: the pockets, the bootleg and the sock are all impossible locations for the cheese.

Semantically, Leena and Raita’s NPs refer to entities that are comparable to Mella’s initial suggestion. Together the participants build up a three-part list (line 1 *sun takin taskussa* ‘in the pocket of your coat’, lines 7 and 16 *saappaanvarressa* ‘in the bootleg’, and line 23 *syylingissä* ‘in the sock’). Sanna closes the sequence by rejecting the suggestions (line 25).

It is noteworthy that with the exception of Henna’s utterance in line 13, which focuses on another ongoing activity, eating, participants in this sequence are engaged in one and the same activity that starts out with Mella’s question in line 1. Through this mutual engagement in the same activity, Mella’s clause remains open for adding new material. I would like to argue here that con-

2. In colloquial Finnish, a negative statement which has the negation verb as its first element can be understood as a question requesting for confirmation even though there is no question clitic (Yli-Vakkuri 1986: 232, Forsberg 1994: 62). The interpretation of the clause as a question is dependent inter alia on prosodic cues: if the primary accent is on the verb, the clause would not be understood as a question but as a statement. With the question clitic, the clause would be interpreted as information seeking, not just seeking for confirmation:

ei-kö ne oo su-n taki-n tasku-ssa.
 NEG+3SG-Q 3PL be 2SG-GEN coat-GEN pocket-INE
 ‘Aren’t they in your coat pocket?’

structing clauses is one kind of conversational activity, and the co-participants Leena and Raita show orientation towards Mella's turn in line 1 by formulating their suggestions in a way that is congruent with Mella's construction. In other words, Leena and Raita's turns are formulated so that they can function as possible additions to Mella's clause.

In conclusion, the NPs on lines 7, 16, and 23 are clausal arguments, because both morphosyntactically and semantically they can be interpreted as parts of the clause in line 1. It is important to note, however, that they have a lot of meaning potential in themselves because they are lexical content words. Their locative case marking also adds to their meaning potential. Thus, they are interpretable independently from the clause on line 1, and the clause is interpretable independently from the locative NPs. In other words, the locative NPs belong to the clause but they do not constitute it in the same way as the core arguments do. They are only slightly more integrated into clausal syntax than free NPs.

Free NPs do not form a part of any clause, but rather form syntactic units of their own, serving various discourse functions (see Chapter 4 for more discussion). In the following example, in line 7, there is an example of an identifying free NP. Prior to this sequence, there has been talk about eating breakfast and whether one is hungry in the morning.

(12) Birthday

- 1 Noora: $\wedge noi$ *ol-i* *ol-lu* *hotelli-aamiaise-lla*.
 those be-PAST+3SG be-PCP hotel-breakfast-ADE
 These guys had breakfast in a hotel (lit. those had been on a
 hotel-breakfast).
- 2 Leena: $\langle F \wedge mm, F \rangle$
 uhhuh
 Yeah.
- 3 Henna: $\wedge j[_1 oo _1]$,
 yeah
 Yeah.
- 4 Sanna: $[_1 \wedge joo _1]$.
 yeah
 Yeah.
- 5 *täy-*, *täyty-y* *men-nä*.
 must-3SG go-INF
 One (I) should go (to eat breakfast in a hotel).

- 6 Noora: *ei-k ol-lu hyvä* [₂⟨X X₂⟩]
 NEG-Q be-PCP good
 It was good, wasn't it?
- 7 Leena: [₂[^]*Klaus* ₂] *Kurki*. ←
 Klaus Kurki (name of a hotel).
- 8 Noora: ... *^joo*.
 yeah
 Yeah.
- 9 Henna: ... *tosi hyvä*.
 real good
 (It was) real good.

In the example, Noora (line 1) is telling Sanna about *noi* 'those' that they have been to eat breakfast in a hotel. The demonstrative pronoun *noi* can be used to refer to co-participants or a co-participant and somebody else (for discussion, see Seppänen 1995, 1998). From the response it seems that Leena and Henna are (among) the people that Noora refers to (cf. the fact that Leena knows the name of the place (line 7) and Henna is able to make an assessment about the quality of the place (line 9). Leena gives a response that projects more on this topic (line 2), and Henna seconds her with her *joo* 'yeah' (line 3). According to Sorjonen (1997), the particle *joo* can be used as a response to show recognition of the referential form used in a prior utterance and, at the same time, to show that continuation is expected. With her turn (lines 4 and 5) Sanna takes the position of one of the recipients of the story Noora has initiated. Noora asks whether it was good (line 6). Her question is formulated so that the question clitic *-k(o)* is attached to the negation verb. In Finnish, when the question clitic is attached to the negation verb *ei* in assessments, the question seeks affirmation (cf. English tag questions, see also fn. 3). Note that the question in line 6 is made without mention of the entity about which the assessment is made. Leena responds to this with the name of the hotel (line 7), which may seem sequentially inappropriate as an answer to the question presented. However, the restaurant Leena names has a good reputation and is known for its good food so that even people who have never visited the place know that they serve good food. Thus Leena responds to Noora's question indirectly.

The response is produced as a free NP that serves an identifying function. The entity being identified, i.e. the hotel, has been mentioned as the first part of a noun compound in line 1, but it has not been directly referred to, since in the NP *hotelliaamiainen* 'hotel breakfast', *hotelli* 'hotel' is general, not specific.

As will be shown in Chapter 4, it is typical of free NPs that they are used to make predications about entities which have not been directly referred to before as such.

I would like to argue that the free NP in line 7 cannot be thought of as being part of an elliptical clause since there is nothing in the context that could have been ellipted. In terms of the speech activities the speakers are engaged in, it does not just expand some previous activity (like the locative NPs in example 11), but rather, it serves a function of its own in identifying the place by naming it. As I will show in Chapter 4, free NPs are a grammatical resource that serves specific functions in discourse. In grammar, they form syntactic units that have equal validity with clauses.

In addition to clauses and free NPs, there is a third type of syntactic unit that can be distinguished, namely particle chains. Particles may form utterances of their own or they may join to form clusters or chains which are constrained by fairly strict ordering principles. For example in English it is possible to say *oh no* as an exclamation, but the reversed *no oh* is not a crystallized expression. The syntax of Finnish particles and particle chains has received little interest in the literature (see however, Sorjonen 1997, A. Hakulinen 1989, 1993), and it falls outside the scope of the present work. However, I would like to point to two principles with respect to which particle chains differ from, say, clausal syntax: first, their syntactic organization seems to be linear and not hierarchical (cf. e.g. verbal government), and secondly, the mutual ordering principles are specific to the individual particles and not general (like “place subject before verb”). In some cases, the particles in a particle chain may stand in a modifier — head relationship where one particle modifies the use of the other particle (for example in *ai joo* the first part *ai* modifies *joo* so that *joo*, which is not usually used as a receipt of prior talk as new information, can be used to display that prior talk has been interpreted as news (Marja-Leena Sorjonen, pc).

I will try to illustrate some of the ordering constraints with the following example. It comes from a conversation between a mother (Tintti) and a daughter (Leena). Before this sequence, the mother has been telling about the kinds of liquor she has purchased for Christmas. She adds one more kind to her list:

(13) Mother and Daughter

- 1 Tintti: *sit mää ost-i-n, m-... vadelma-mansikka-likööri-i.*
 then I buy-PAST-1SG raspberry-strawberry-liqueur-PTV
 Then I bought, stra-... raspberry-strawberry liqueur.

- | | | | |
|---|---------|--|---|
| 2 | Leena: | <i>aha.</i> | ← |
| | | Uhhuh. | |
| 3 | | ...(0.5) <i>jaa</i> ^ <i>jaa</i> . | ← |
| | | yeah yeah | |
| | | I see, I see. | |
| 4 | | ...(0.7) ^ <i>no niin</i> . | ← |
| | | well so | |
| | | That's it. | |
| 5 | Tintti: | <i>ne on</i> , | |
| | | they be-3SG | |
| | | They are, | |
| 6 | Leena: | ^ <i>suur-hankinna-t</i> . | |
| | | big-investment-PL | |
| | | big investments. | |
| 7 | Tintti: | <i>ne on</i> ^ <i>sievä-t tuol kirja-hylly-n pää-l</i> . | |
| | | they be+3SG cute-PL there book-shelf-GEN top-ADE | |
| | | They are cute there on top of the book shelf. | |

In example (13), there are three intonation units (lines 2–4) that consist solely of particles. Each of these units serves a special purpose here: the *aha* in line 2 acts as a mere receipt of the foregoing talk, whereas the *jaa jaa* in line 3, which is uttered after a considerable pause of 0.5 seconds, signals closing of the topic, namely the list of purchases. As Tintti does not continue, Leena closes the sequence with *no niin* in line 4. According to Raevaara (1989: 149), the particle chain *no niin* can be used to terminate a sequence in conversation. In line 5, Tintti shows acceptance of Leena's interpretation that the topic has been closed by starting with a nominal clause construction. In line 6, Leena breaks in and completes the nominal clause started off by Tintti with an NP *suurhankinnat* 'big investments'. However, Tintti does not quite accept this but recycles the whole construction in line 7. It is noteworthy that all three particle constructions form intonation units of their own ending in a final intonation contour. Each displays a certain kind of understanding of the previous talk. In line 4, if the particles *no* and *niin* were uttered in reverse order we would have an acceptable utterance but the interpretation would be totally different: *niin no* would project some kind of re-evaluation of the described sequence of events. Lines 5–6 would not be appropriate continuations after *niin no*. Thus, a change in the ordering of particles causes a change in the projection for the following talk.

2.1.3 Summary

In this section, I have argued for a view which sees syntax as emergent structures. With several examples, I have tried to show that the interpretation as to what forms a clause depends on the activity that the participants are engaged in.

I have further shown that clauses are not the only kind of syntacticized units that need to be taken into account when describing grammar from the viewpoint of spoken discourse. Tao (1996) shows that the clause may not emerge equally clearly as a level of syntactic organization in all languages. Based on evidence from Mandarin spoken discourse, he argues that phrasal organization is much more crucial to the understanding of Mandarin syntax than the clause level. This provides further evidence for my claim that we need to reconsider the notion of clause and its centrality in syntactic description. We are only starting to build a cross-linguistic understanding of what clauses are as discourse constructs.

I have shown in this section that free NPs and particle chains form syntacticized units which are not reducible to clauses. I would like to further emphasize that free NPs, particle chains and clauses may not always be clearly distinguishable. However, if we think of syntactic constructions as instantiations of abstract constructional schemas (see Section 1.1), it should be no problem if some instantiations do not correspond exactly to any schema or do not realize every part of the schema.

In what follows, I will discuss the construction of clauses in conversation. In my view, clausal syntactic relations emerge most clearly in the relationship between the verb and its core arguments. In the following sections I will show how these relations are coded and how core arguments are distinguished from obliques.

2.2 Coding strategies for syntactic functions

In the typological literature, a distinction is often made between three basic coding strategies for syntactic functions: nominal case marking, cross-referencing or agreement, and word order (see e.g. Andrews 1985: 71, Croft 1991). The three coding strategies are all employed in Finnish, but none of them, taken by themselves, resolves the question of how syntactic functions are distinguished in the clause core or how core arguments are distinguished from obliques. My aim with this discussion is to show how clausal syntactic relations emerge as a result of the interaction of these coding strategies.

Across languages, it is quite common for personal pronouns to have a coding pattern that is distinct from that of other pronouns and full NPs. For example, in English, personal pronouns realize a nominative-accusative marking system, whereas full NPs have no case marking. There is a rich literature on splits in coding systems, especially concerning the ways in which they manifest themselves in case marking and agreement patterns (see e.g. Silverstein 1976, Durie 1988, Mithun 1991).

In his seminal article on ergative splits, Silverstein (1976) proposed that the splits are not random, but rather, correspond to a hierarchy of “inherent lexical content” of noun phrases. In the hierarchy, the inherent referential content of noun phrases is analyzed in terms of binary features (such as $\pm$ ego, $\pm$ tu, $\pm$ plural etc.). Personal pronouns rank highest in the hierarchy, and they are followed by third person pronouns, proper nouns and kin terms, then nouns that refer to humans, then animates, and inanimates (Silverstein 1976:122). According to Silverstein (1976:113), this hierarchy expresses “the semantic naturalness for a lexically-specified noun phrase to function as agent of a true transitive verb, and inversely the naturalness of functioning as patient of such”.

According to Silverstein, the actual morphosyntactic case marking configurations are not only affected by the interaction of inherent referential content with syntactic roles (or, in Silverstein’s terminology, “case relations” such as “agent-of” and “patient-of”), but also by two additional variables, namely clause linkage types, and reference-maintenance relations of arguments (Silverstein 1981:228–231). Thus, coding patterns are also influenced by discourse factors. Later research on topic continuity and information flow (see e.g. articles in Givón 1983, Du Bois 1985, 1987a) can be seen as developing Silverstein’s proposal in suggesting possible discourse factors which explain the emergence of split patterns.

The fact that personal pronouns often show a distinct coding pattern has led researchers to exclude them from the description of alignment systems (cf. e.g. Siewierska 1997). In my view, one of the most interesting — but unfortunately rarely discussed — insights in Silverstein’s work is that personal pronouns are the heart of coding systems and that the functioning of coding systems, especially those showing split patterns, cannot in fact be properly understood without taking into account personal pronouns. Therefore, in the discussion of case marking and other coding strategies which follows, personal pronouns are treated as central and not as peripheral to the system.

I will first discuss the role of case marking in the coding of syntactic functions. Then I will move on to agreement and word order. Finally, I will

summarize the discussion on the three coding strategies and look at the concept of clause again in the light of this discussion.

2.3 Case marking: General principles

In his classic discussions of case marking in the Slavic languages, Jakobson (1971 [1936], 1984 [1958]) distinguishes the nominative and the accusative as direct cases which “amount to an opposition of a case signalling the object of a process to a case not containing this signalization” (1984 [1958]: 110), in other words, the nominative and the accusative code participants of a (transitive) process. The direct cases are opposed to the others, termed oblique.

Jakobson’s distinction between direct and oblique cases reflects the mainstream structuralist position which, according to Nichols (1983), maintains that cases can be divided into two sets, those that are more abstract and mark syntactic relations (core cases) and those that are more semantic (obliques). Nichols (1983: 170) shows that the distinction between core and oblique roles evolved out of a distinction based on case forms, namely the classical distinction between nominative and oblique case forms, where the latter referred to all non-nominative cases. This distinction was combined with a division of cases into those that are due to rection (verbal government) vs. others (see also T. Itkonen 1979:97).

Foley and Van Valin maintain (1984:79), “The main coding feature distinguishing core from peripheral arguments is that core arguments tend to occur in unmarked morphological or syntactic forms and peripheral arguments in marked, often adpositional, codings.” Some languages show the distinction between core and oblique quite clearly in their case marking: for example in Obolo (Benue-Congo), NPs in the core roles do not have case marking, whereas oblique NPs are marked as such with the help of prepositions. But even in Obolo, there are exceptions to this rule; *inter alia*, certain time and measurement adverbials may appear without a preposition — as if they were core NPs — to indicate aspectual distinctions, such as more complete coverage or more thorough effect of the process described by the verb (Helasvuo 1992).

From the perspective of the grammaticization, the distinction between core and oblique cases does not seem as clear-cut. There is plenty of evidence in the literature that there is a universal tendency for locative elements to develop more grammatical meanings (see e.g. Heine et al. 1991:113–118; Harris & Campbell 1995:339–341). Thus, in the light of grammaticization

studies, the case system appears not so much as a set of Jakobsonian binary oppositions, but rather, a dynamic system that is subject to change. By studying spoken discourse, we may get a better understanding of how the system evolves and what kind of functions it is designed to serve. Interpreting Jakobson's definition of core and oblique from this perspective, we could say that the core cases serve to code the major participants in a process whereas the oblique cases fail to code them.

In my study, I have distinguished between the core and oblique cases in a very simplistic way on the basis of what syntactic functions the NPs are used for in discourse in order to determine whether the two sets of cases show different discourse profiles. As we will see below, they indeed do. However, the study also reveals that within the two sets there are cases which show characteristics of both core and oblique cases, namely the partitive and the adessive. The case alternation of the core cases is very complex, and there is a rich literature on its intricacies. In the discussion below, I will try to approach the case system from the viewpoint of its use in discourse. My main points will be: (i) in the case system the nominative stands out in many respects as different from the other cases, and (ii) the case alternation of the grammatical cases is easier to understand if the marking of personal pronouns is taken as a starting point. After a summary on the interplay between the grammatical cases, I will move on to the obliques.

2.4 Case marking in Finnish

Finnish is well known for its rich case system, and it has been discussed extensively in the literature, especially in terms of the structural oppositions that the grammatical cases take part in (see e.g. Siro 1964, T. Itkonen 1979, Heinämäki 1983, Leino 1990a). As a basis for further discussion, I will give an overview of the case system in this section. I will first discuss the system from a traditional structural point of view and then I will outline an alternative approach.

Table 1 gives an overview of the system with examples of the singular forms.³ Most of these cases also inflect in the plural, except for the accusative, which does not have a plural form. Note that personal pronouns have a special accusative form (ending *-t*), and that they also inflect in the plural accusative

3. Table 1 has been modelled after Shore (1992). The translation equivalents given to the example words are hers.

(although this is not shown in the table). Only the most productive cases are included in the table.⁴

In Table 1, the cases have been divided into two groups, the grammatical cases (upper half of the table, marked with Roman I) and the obliques (lower half, marked with Roman II).⁵ While the oblique cases have more independent semantic meaning potential and are constrained by semantic factors in their use, the grammatical cases have been grammaticized to a greater extent, i.e. they are constrained by syntactic factors (e.g. negative polarity, aspect) in their use.

Table 1. The Finnish case system

	Case form	Case ending	Example	Translation
I	Nominative	-Ø	<i>talo</i>	a/the house
	Accusative	-n	<i>talon</i>	a/the house
	Acc of pers.pron.	-t	<i>minut</i>	me
	Partitive	-(t)A	<i>taloa</i>	(of) a/the house
	Genitive	-n	<i>talon</i>	of a/the house
II	Essive	-nA	<i>talona</i>	as/for a/the house
	Translative	-ksI	<i>taloksi</i>	into (a/the) house
	Inessive	-ssA	<i>talossa</i>	in(side) the house
	Elicative	-stA	<i>talosta</i>	from inside the house
	Illative	-Vn,-hVn,-seen	<i>taloon</i>	into (a/the) house
	Adessive	-llA	<i>talolla</i>	by/on/near a/the house
	Ablative	-ltA	<i>talolta</i>	from the house
	Allative	-lle	<i>talolle</i>	to the house

Traditionally, the Finnish cases have been defined on morphological grounds, except for the nominative and the accusative, which have been defined on syntactic as well as morphological grounds. The accusative is often described as a “syntactic case” (see Setälä 1908 [1898]: 51–52; the literature thereafter has followed this tradition; see e.g. Penttilä 1963, Hakulinen & Karlsson 1979),

4. Of the remaining cases, the instructive is not productive any more (e.g. *käsi-n* ‘by hand’), whereas the abessive is productively used to form infinitival constructions (e.g. *odotta-ma-tta* ‘wait-INF-ABE, without waiting’). The comitative (e.g. *vaimo-ine-en* ‘wife-COM-POSS.SUFF, with his wife’) is not used very much at all, but instead, an adpositional phrase formed with the postposition *kanssa* ‘with’ is being used to convey roughly the same meaning (e.g. *vaimonsa kanssa* ‘with his wife’).

5. The genitive is somewhat problematic here (for discussion, see Helasvuo 1996b and Section 2.4.2 below).

which is in opposition to the partitive in the object role (for an opposing view, see Shore 1992, Nemvalts 1994, who reject the term accusative altogether and use genitive for all *-n*-marked cases).⁶ Two kinds of accusative markers are distinguished, namely the so-called “genitive-accusative” with the marker *-n*, and the so-called “nominative-accusative” that has no marker. What is called the nominative in this view is residue: it is a case with a zero marker for NPs that are not in the object role.

I find this usage confusing and will not adopt it here. I prefer to use the term nominative for zero marking irrespective of what the syntactic function of the NP is. In the nominative plural, there is no case marking and *-t* marks plurality.⁷ Accordingly, I reserve the term accusative for *-n*-marked NPs. The accusative is syncretic with the genitive, which has the same marker. However, they differ in that the accusative does not inflect in the plural, whereas all genitive NPs do. The two cases were historically distinct (see below Section 2.4.2 for discussion), and in the present-day language, they are easily distinguishable because they are used in different syntactic contexts and because they differ in the discourse functions they serve (see Section 2.4.2).

Table 1 looks at the case system from a structural point of view, as a set of choices available for the speaker. In this study, I take an alternative perspective: instead of describing the case system in terms of the structural oppositions the cases take part in, I discuss the system as a discourse construct. In this approach, discourse tokens, instead of types, form a central focus of interest. The system behind the tokens will be approached from the viewpoint of the different functions the cases serve in discourse. My hope is that this will provide us a way to understand why the system is the way it is.

Despite the rich set of choices the case system offers, the use of the different cases clusters around only a few cases. In Table 2 below, the cases are organized according to discourse frequency.⁸

6. There is some ambivalence in the contemporary literature on object case marking in Finnish: it is admitted that the nominative is an object-marking case on a par with the accusative and the partitive, but at the same time, the case alternation is described as a structural opposition between the partitive and the accusative, the latter of which encompasses both the nominative and the (morphological) accusative (see Leino 1991, Vilkkuna 1996).

7. The same *t*-ending is used in the 3rd person marking on the verb (*-va + t*) to code plurality.

8. Note that Table 2 counts noun phrases, not nouns. In Finnish, modifiers almost always

Table 2. Case marking of NPs: discourse frequencies.

Case of NP	N	%
Nominative	967	55.0
Partitive	266	14.8
Adessive	108	6.0
Inessive	106	5.9
Accusative	69	3.8
Elative	63	3.5
Illative	60	3.3
Genitive	48	2.7
Allative	41	2.3
Essive	18	1.0
Translative	17	0.9
Ablative	11	0.6
Total	1794	100

Table 2 shows that the nominative (i.e. the zero-marked case) overwhelmingly outnumbers all other cases. Räsänen (1972) reports that in his dialectological data from interviews, around 35% of the nouns were in the nominative. Note that Räsänen counts every nominal element separately, whereas I have based my counts on noun phrases (see previous footnote). In Räsänen's study all NPs with modifiers are counted as containing several case-marked nominals, whereas in my study, modifiers are not counted separately. Previous studies have shown that the nominative is the most frequent case in written discourse also, although the difference is not so striking (in the written corpus of the Finnish syntax archive, a little less than 30% of nouns were in the nominative; Karlsson 1982:308). Only the partitive comes anywhere close to the nominative with its almost 15%. All other cases seem quite marginal in terms of discourse frequencies. Of course, Table 2.2 gives only raw frequencies that need

agree with their head in case and number. For example, if the head is in the inessive *talo-ssa* 'in the/a house', so is the modifier, e.g. *piene-ssä talo-ssa* 'in the/a small house'. Possessives are exceptional here, thus we have *Miko-n talo-ssa* 'in Mikko's house'. Therefore, Table 2 gives a somewhat biased view of the frequency of the genitive, because genitives modifying NPs were not counted, but only genitives that form an NP of their own, e.g. *minun* in the following example:

Minu-n täyty-y lähte-ä.
 1SG-GEN must-3SG go-INF
 I must go.

to be interpreted in the light of the discourse needs that the cases serve. This is what I will attempt to do in the following subsections.

In what follows, I will discuss the so-called grammatical cases — the nominative, the accusative, and the partitive — that are used to mark the syntactic functions of the clause core. We will look at their respective discourse profiles. After a summary I will complete the discussion of case marking by discussing the oblique cases.

2.4.1 Nominative

The high frequency of the nominative observed in Table 2 cries for explanation: Why is this one case so popular? What kind of work does it do in discourse? What is its relationship to other cases?

Unlike many other cases, the nominative can accommodate almost any syntactic function. Consider the following examples (N.B. the nominative is not glossed in the morphological gloss line!):

(14) Mother and Daughter

Leena: *mä jo-isi-n joka päivä*
1SG drink-COND-1SG every day
I would drink (alcohol) every day.

(15) Women

Eila: *teh-t-i-i semmonen,... (1.3) semmonen vuoristomatka,*
do-PSS-PAST-PERS a-kind-of a-kind-of mountain-trip
We made a kind of a mountain trip. (Lit. “A mountain trip was made.”)

The nominative can be used to mark the subject, as in example (14) *mä* ‘I’. In example (14) there is another nominative NP, namely the time adverbial *joka päivä* ‘every day’. However, the use of the nominative in time or measurement adverbials is semantically restricted to NPs in which the head noun carries a clear meaning of measuring either time or distance (such as hour, day, kilometer). In example (15) there is a passive clause in which the object *semmonen vuoristomatka* ‘a kind of a mountain trip’ stands in the nominative (for a discussion of the analysis of the object in passive clauses in Finnish, see Section 1.3 above). The nominative can also be used in predicate nominals (see example 16) and free (unattached) NPs (17).

(16) Birthday

Sanna: *ne oli aika iso-t pala-t.*
 3PL were quite big-PL piece-PL
 They were pretty big pieces.

In example (16) both the subject *ne* ‘they’ and the predicate nominal *aika isot palat* ‘pretty big pieces’ stand in the nominative. Example (17) illustrates the use of the nominative in free NPs.

(17) Birthday

- 1 Sanna: *vienosti ilmoitan että [voin kuvitella tähän,]*
 shyly announce-1SG that can-1SG imagine here
 I shyly announce that [I can imagine here]
- 2 Henna: [*(X) herkku-juusto-a.*]
 delicacy-cheese-PTV
 [delicious cheese.]
- 3 Sanna: *brie-n,*
 brie-ACC
 the Brie,
- 4 ... *näin,*
 like this,
- 5 Henna: @@
- 6 Sanna: *ja sitten voi-mme kuvitella näin Masurka-n.*
 and then can-1PL imagine so M.-ACC
 and then we can imagine Masurkka (name of a cheese) like this.
- 7 Henna: *taas nää keisari-n uude-t vaattee-t.* ←
 again these emperor-GEN new-PL cloth-PL
 the emperor’s new clothes again.
- 8 *aina nää Sanna-n hankinna-t.* ←
 always these S.-GEN purchase-PL
 always these purchases by Sanna.

In example (17), the participants are having dinner together. Sanna was supposed to bring cheese to the party, but she didn’t (cf. ex. 10 above). Nevertheless she introduces the non-existent cheese to the other participants ‘we can imagine brie and Masurkka here’ (lines 1, 3, and 6). Henna responds to this pretense play by laughing (line 5) and then describes Sanna’s activities with a free NP *nää keisarin uudet vaatteet* ‘these new clothes of the emperor’ that

alluding to the fairytale where the emperor goes around naked but all his loyal servants pretend that he is wearing new clothes. Line 8 contains yet another free NP *nää Sannan hankinnat* ‘these purchases of Sanna’, also in the nominative.

Table 3 illustrates the distribution of nominative NPs across syntactic roles.

Table 3. Distribution of nominative NPs across syntactic roles.

Syntactic function	N	%
Subject	620	62.9
Object	50	5.1
Free NP	104	10.5
E-NP	79	6.0
Predicate nominal	61	6.2
Other	72	7.3
Total	986	100

As can be seen from Table 3, most nominative NPs function in the clause core as subjects, objects, E-NPs and predicate nominals, in addition to being used as free NPs. No other case is used in such a wide variety of syntactic functions. The adaptability of nominative NPs for such a wide range of functions explains why nominatives are so frequent in the data, compared to the other cases, as shown in Table 2. Free NPs, of course, are primarily characteristic of spoken discourse, which helps partly to explain the large number of nominatives found in my study, compared to earlier studies which were based on written data (Karlsson 1982). In addition, it has been noted that spoken clauses do not accommodate as many oblique arguments as written clauses (Chafe 1979:164 for English, Pajunen and Palomäki 1985:7 for Finnish).

Our discussion so far concerned NPs that stand in the nominative without making any finer distinctions. The picture changes dramatically if we take into account what kind of NP is in question. I will illustrate this with Tables 4–6. For personal pronouns, the nominative is a true subject marking case, and the marking of the core roles follows a canonical nominative-accusative coding type where the subject is always in the nominative and the object is marked either with the accusative or the partitive (Table 4). If a personal pronoun is in the nominative, it will always function as a subject in the clause. Consider example (18).

(18) Mother and daughter

Tintti: *jos sää vo-isi-t se-n teh-dä,*
 if 2SG can-COND-2SG 3SG-ACC do-INF
 if you could do that,
niin mä o-isi-n tosi ilone.
 then 1SG be-COND-1SG real glad
 then I would be real glad.

In example (18), the 2nd person pronoun *sää* ‘you’ is in the nominative and functions as the subject. The 3rd person pronoun *se* ‘it’, functioning as the object of the if-clause, stands in the accusative. The then-clause is a predicate nominal clause where the 1st person pronoun *mä* ‘I’ functions as the subject. Table 4 illustrates case marking in the core roles for personal pronouns.

Table 4. Case marking in the core roles. Personal pronouns.

Subject	Object
nominative	accusative or partitive

The case marking patterns are slightly different if we look at singular full NPs and pronouns other than the personal ones. Table 5 serves to illustrate this. As Table 5 shows, the nominative is no longer a subject marking case, but instead, it is used for subjects and, if there is no nominative subject in the clause, for objects as well. But if there is no nominative NP subject in the clause, and therefore, there is no need to distinguish the object from the subject, the object can be in the nominative.⁹

Consider example (19) which provides an example of a clause without a nominative subject.

(19) Mother and daughter

Tintti: *pudote-tt-i-in niin helveti-n upee kuvasto.*
 drop-PSS-PAST-PERS SO hell-GEN magnificent catalog
 They dropped a damn beautiful catalog (through the mail) (lit.
 was dropped so damn beautiful catalog).

9. As can be noted in Table 3, only around 5% of nominative NPs function in the object role.

Table 5. Case marking in the core roles. Other singular pronouns and singular full NPs.

Singular	
Subject	Object
nominative (no subject)	accusative or partitive nominative or partitive

In example (19), the object NP *helvetin upee kuvasto* ‘a damn beautiful catalog’ is in the nominative because there is no nominative subject in the clause. There is a principle of economy at work here: the object is marked with the accusative only if there is a need to distinguish it from the subject (cf. Du Bois 1985). If there is no nominative subject, the object can take the nominative case. (For further discussion, see Helasvuo 1997.)

To summarize, the marking of other pronouns and singular full NPs does not follow a canonical nominative–accusative pattern. Instead, the nominative functions more like an absolutive.

Table 6 illustrates case marking of full plural NPs and plural pronouns (except for personal pronouns). It shows that the marking exhibits a neutral marking system where case marking does not distinguish between the subject and object roles.

Table 6. Case marking of the core roles. Plural full NPs and plural pronouns (except for personal pronouns).

Plural	
Subject	Object
nominative	nominative or partitive

As Table 6 shows, clauses where both the subject and the object are plural full NPs and plural pronouns (not personal pronouns though!) allow for both the subject and the object to be in the nominative. Example (20) illustrates this.

- (20) (From Palander 1991:242)
- Ei*

NEG+3SG
- tarvihitte-is*

need-COND
- tuola työ-mua-la*

there work-site-ADE
- ennee*

anymore
- pui-hen*

tree-PL+GEN
- kansap*

with
- punnerto-o.*

wrestle-INF
- One wouldn’t need to wrestle with logs anymore on the site.

Raskaa-t työt tek-is konnee-t. ←
 hard-PL job-PL do-COND+3SG machine-PL
 The hard work would be done by machines (lit. Machines would do the hard jobs.)

In the latter clause in example (20), both the subject *konneet* ‘machines’ and the object *raskaat työt* ‘hard jobs’ are full NPs. The word order is quite unusual, as the object is preverbal and the subject postverbal (for further discussion, see Section 2.6 below). In fact, the only way to determine the grammatical roles in this clause is to base the analysis on our world knowledge. It is important to note here that clauses where both the subject and the object are plural full NPs or plural pronouns are rare in natural discourse.¹⁰

Interestingly enough, cross-linguistic studies have shown that there is a clear tendency in spoken discourse not to have more than one full NP argument per clause (see e.g. Du Bois 1985 on Sakapulteko, Durie 1988 on Acehnese, Lambrecht 1987 on French, Ochs 1988 on Samoan, Kärkkäinen 1996 on English). This full NP argument tends to function as the object or intransitive subject but is rarely the transitive subject (Du Bois 1985). Thus, the need to distinguish between a full NP subject and a full NP object does not arise very often. Even more rarely does one find cases where they would both be in the plural. In my data, there were no instances where both the subject and the object would have been plural full NPs (see Du Bois 1987b:215 on the avoidance of “double positives”). This, of course, increases the range of possibilities of using nominative to mark both subjects and objects.

To summarize, I have tried to show in this section that the nominative is a multifaceted and multifunctional case. In the marking of personal pronouns the use of the nominative follows a classical nominative–accusative coding pattern. In the marking of other pronouns and singular full NPs the nominative functions more like an absolutive. The marking of plural pronouns and plural full NPs resembles a neutral marking pattern.

10. Palander (1991) provides some examples of clauses with two plural NP arguments. Her data consists of interviews collected for dialectological study. From her tables it can be inferred that clauses with two plural NP arguments made up around 3.7% of the clauses in her data (Palander 1991:246–247).

2.4.2 Accusative

In the literature on case marking, the accusative is usually defined as an object marking case that stands in opposition to the nominative. In languages with a nominative–accusative marking system, both transitive and intransitive subjects are marked with the nominative, whereas objects are marked with the accusative. As we saw in the previous section, Finnish shows nominative–accusative alignment in its coding system, although the system is not quite so straightforward. In what follows, I will first discuss the marking of personal pronouns and compare it to the marking of other pronouns and nouns. Then I will discuss the history of accusative marking and contrast it with the genitive.

In personal pronouns in Finnish, the accusative alternates with the partitive as an object marking case (cf. Table 7; in the table, the inflection of the 1st person pronouns *minä/me* ‘I/we’ is given as an example). (Recall that personal pronouns are always in the nominative when functioning as subjects.)

Table 7. Object marking in Finnish. Personal pronouns.

	Singular		Plural	
	PTV	ACC	PTV	ACC
Personal pronouns	-A <i>minua</i>	-t <i>minut</i>	-tä <i>meitä</i>	-t <i>meidät</i>

In Finnish linguistics, the case alternation in the marking of the object is often discussed with respect to three parameters, namely (i) positive vs. negative polarity of the clause, (ii) aspect, and (iii) definiteness vs. indefiniteness of the object NP (see e.g. Vilkuna 1996). I will now discuss these parameters in the light of examples.

If the clause expresses negative polarity (or even if it is positive in form but has a negative interpretation), the object stands in the partitive (ex. 21a and 21b, which is a fabricated example).

- (21) a. Alko
 Aarre: *e-n mä oo huoman-nu ero-o,*
 NEG-1SG 1SG be+INF notice-PCP difference-PTV
 I haven’t noticed any difference.

b. (Fabricated example)

mä oo-n huomana-^{nu} ero-n.

1SG be-1SG notice-PCP difference-ACC

I have noticed a difference.

In example (21a), the object NP *eroo* ‘difference-PTV’ stands in the partitive, because the clause is negative, whereas in (21b) the object *eron* ‘difference-ACC’ is in the accusative as the clause expresses positive polarity.

Likewise, if the clause is unbounded aspectually, the object has to be in the partitive. According to Heinämäki (1983, 1994), accusative marking can be used to set a bound to the activity or state described by the verb, thus making it telic, whereas the partitive sets no such bound (ex. 22).

(22) Women

Linnea: *hän ... sirpi-n pist-i to-ta naru-n pää-hän,*
 3SG sickle-ACC put-PAST+3SG that-PTV rope-GEN end-ILL
 He put a sickle at the end of a rope

ja hän veti pitkin pohja-a si-tä sitte perä-ssä.
 and 3SG pulled along bottom-PTV it-PTV then behind-INE
 and then pulled it along the bottom (of the sea) behind (himself) (in order to cut weeds).

In example (22) line 1, accusative marking is used to set a bound to the activity, thus resulting in *sirpin* ‘sickle-ACC’, whereas no such bound is set in the clause in line 2, and accordingly, the object *sitä* ‘it-PTV’ is in the partitive.¹¹ The action described in the first clause is punctual, and therefore, the object is in the accusative, whereas the second clause describes a process with no distinct endpoint, and accordingly, the object is in the partitive.

Finally, if the referent of the object NP is uncountable and indefinite, it stands in the partitive (ex. 23a). Otherwise, it is in the accusative (ex. 23b which is a fabricated example).

11. *To-ta* ‘that-PTV’ is a partitive form of the demonstrative ‘that’. In this example, however, *tota* is being used as a particle: it appears in a crystallized form (the partitive) irrespective of the syntax of the rest of the clause. There is no in-depth study of the use of *tota* as a particle. It could be described very roughly as a marker of hesitation.

- (23) a. Mother and Daughter
 Tintti: *ai sää juo-t* *Cointreau-t*,
 oh 2SG drink-2SG Cointreau-PTV
 Oh you are drinking Cointreau.
- b. (Fabricated example)
 jo-i-t *sä jo* *se-n* *Cointreau-n*.
 drink-PAST-2SG 2SG already it-ACC Cointreau-ACC
 Did you already drink the Cointreau (a glass/a bottle)?

In (23a) the object *Cointreau-t* ‘Cointreau-PTV’ is interpreted as indefinite with the partitive, whereas in (23b) the object NP *sen Cointreaun* ‘the Cointreau-ACC’ is marked as definite and countable (meaning a portion of the liqueur) with the help of the accusative.

We have seen that in the marking of personal pronouns, the Finnish case system conforms to a classical nominative-accusative system: subjects, both transitive and intransitive, stand in the unmarked form, the nominative, and objects are always marked either with the accusative or the partitive.

The marking system is more complicated for pronouns other than the personal ones and for nouns. Consider Table 8. (Grey areas indicate markings that are not available in a given context, e.g. nominative is not available for objects in the singular if there is an NP subject.)

Table 8. Object marking in Finnish. Full NPs and pronouns other than personal ones.

	Singular			Plural	
	PTV	ACC	NOM	PTV	NOM
-in clauses with an NP subject	-(t)A	-n		-(t)A	-Ø
-in clauses without an NP subject	-(t)A		-Ø	-(t)A	-Ø

In addition to the three parameters mentioned above (negative polarity, aspect, and indefiniteness), the marking of full NPs and pronouns other than the personal ones is sensitive to whether there is (or could be) an NP subject in the clause. If there is no subject, then the object can take the nominative (ex. 24; consider also example 19). In the plural, the accusative is not available as a case marking alternative for the object: if none of the three parameters enforces the partitive, the object stands in the nominative (ex. 25).

- (24) (Fabricated example)

sirppi pistet-t-i-in naru-n pää-hän.

sickle put-PSS-PAST-PERS rope-GEN end-ILL

A sickle was put in the end of a rope (or: they put a sickle in the end of a rope).

- (25) (Fabricated example)

hän pist-i sirpi-t naru-n pää-hän.

3SG put-PAST+3SG sickle-PL rope-GEN end-ILL

He put the sickles in the end of a rope.

In example (24) the object NP *sirppi* ‘sickle’ stands in the nominative because there is no subject NP in the clause.¹² In example (25) the object *sirpit* ‘sickles’ is in the plural, there is no reason to use the partitive, and therefore it is in the nominative.¹³

Historical studies have shown that in Proto-Uralic the partitive was a pure local case with a separative meaning (for more discussion, see the next section). In the marking of the object, the accusative (ending **-m*) alternated with the nominative so that definite singular objects were marked with the accusative, while all other objects were unmarked, i.e. in the nominative (see e.g. Setälä 1880, E. Itkonen 1972: 183). As the use of the partitive expanded, the semantic opposition between the nominative and the accusative based on definiteness was shaken, and the object marking system underwent a restructuring process. In Proto-Finnic, there were interesting phonological changes which are very likely to have had an impact on the restructuring process. Namely, word-final *-m* became *-n*. This change made the former accusative ending *-m* coalesce with the genitive ending *-n*. After this change, the accusative and the genitive could no longer be distinguished on the basis of their morphological form; only syntactic distribution keeps them apart (in principle, the accusative is a marker of verbal arguments, whereas the genitive marks adnominal modifiers and objects of adpositions).

In the western dialects of Finnish, the *-n*-marker was used for personal pronouns as well as for full NPs in the marking of the accusative. In the eastern dialects, however, a special accusative marker *-t* was used for the marking of

12. In Finnish, the passive does not function to change the analysis of the object (for discussion, see Section 1.3.5).

13. The *t*-ending in *sirpi-t* ‘sickles’ is a plural marker, not a case ending. The same plural marker is used in the verbal person marking in the 3rd person plural.

personal pronouns, thus making the accusative and the genitive easily distinguishable. The *-t*-marker was adopted to standard written Finnish in the 19th century, and now it is widely spread in spoken Finnish. (Häkkinen 1994: 197.) For example in my data, there are no examples of the *-n*-accusative in the marking of personal pronouns and the *-t*-marker is used exclusively. This development is interesting in light of our discussion of marking systems: the use of the special accusative marker made it possible to avoid the accusative/genitive syncretism which was perhaps confusing in personal pronouns. According to Silverstein, personal pronouns (especially pronouns referring to the speech act participants) function least naturally as objects (patients of transitive verbs), and inversely, most naturally as subjects (agents of transitive verbs). If we look at the distribution of 1st and 2nd person pronouns across grammatical role, we can note that only 4 out of 297 object NPs are 1st or 2nd person pronouns. Thus, 1st and 2nd person pronouns functioning as objects are marked both from a semantic point of view and from a discourse perspective. It is no wonder then that they are always morphologically marked when functioning as objects.

In recent years, several scholars in Finnish linguistics have suggested that there are no grounds for distinguishing the accusative case from the genitive in the modern language, but instead, both should be called the genitive (see e.g. Shore 1992, Vilkuna 1996). In my view, however, this analysis makes the description of the genitive unnecessarily complicated, and therefore, I will not adopt it here.

Given the fact that personal pronouns functioning as objects are so rare in discourse, it may sound odd to claim that the marking of personal pronouns is crucial to the understanding of coding systems, including object marking. However, the obligatory marking of 1st and 2nd person pronouns when functioning as objects also works the other way around: a 1st or 2nd person pronoun that stands in the unmarked form, the nominative, cannot function as an object but has to be the subject. In some sense, only the nominative of personal pronouns is a true nominative that codes the subject, whereas the nominative of other pronouns and full NPs is more like an absolutive, and as such, not distinctively a subject marking case. As I will show in Chapter 3, 1st and 2nd person pronouns play a central role in the grammaticization of the subject role.

2.4.3 Partitive

The partitive originates in an old separative case, but in the modern language it is mainly used as a grammatical case (for a discussion of this grammaticization process, see Helasvuo 1996b). In an interesting way, the effects of this grammaticization process can still be observed in the discourse use of the partitive in the present-day data. Namely, the discourse profile of the partitive has features that are more characteristic of the oblique cases than the grammatical ones. In this section, I will discuss the different functions that partitive NPs serve in discourse and the different factors that constrain the use of the partitive. The common denominator in the various usages is low transitivity.

Almost half of the 266 partitive NPs in my data function as objects (see ex. 26). Others serve as predicate nominals (27), free NPs (28), or even as intransitive subjects (see below Chapter 3). Some function as the core NP in presentational constructions (the E-NP, see Chapter 3). Consider the following examples:

(26) Mother and Daughter

Leena: *mä oo-n niin paljon syö-ny vitamiine-i,*
 1SG be-1SG so much eat-PCP vitamin-PL+PTV
 I have eaten so many vitamins.

(27) Brothers

Simo: *ne on ol-lu iha,... sopus-i-a veikko-j-a*
 3PL be+3SG be-PCP quite friendly-PL-PTV guy-PL-PTV
 They have really been friendly guys.

(28) Birthday

- 1 Sanna: *jo-t-ka mä jemmaa-n ja anna-n aamu-lla.*
 REL-PL-REL 1SG hide-1SG and give-1SG morning-ADE
 (Flowers) that I would hide and give in the morning.
- 2 *mut ei nyt sii-hen aika-an illa-sta*
 but NEG+3SG now 3SG-ILL time-ILL night-ELA
 But at that time of the night you can't
- 3 *[mi-stä-ä sa-a].*
 nothing-ELA-PRON get-3SG
 get (flowers) anywhere.
- 4 Henna: *[voi niin kau]ni-i-ta ajatuks-i-a.* ←
 oh so beautiful-PL-PTV thought-PL-PTV
 Oh such beautiful thoughts!

In example (26) the object *vitamiinei* ‘vitamins-PTV’ is in the partitive. In this case, the use of the partitive signals open quantification of the object referent. In (27) the partitive NP *sopusia veikkoja* ‘friendly-PTV guys-PTV’ functions as a predicate nominal, and in (28) line 4, the NP *kauniita ajatuksia* ‘beautiful thoughts-PTV’ serves as a free NP (see Chapter 4).

Transitivity has traditionally been considered a feature of the verb: a transitive verb takes two arguments, an agent (subject) and a patient (object). From this viewpoint, clauses with partitive objects are no different from clauses with nominative or accusative objects in terms of transitivity. As an alternative to this view, Hopper and Thompson (1980) have proposed that transitivity should be considered a feature of the clause as a whole rather than a characteristic of a given verb. Their scalar notion of transitivity turns out to be useful for understanding the use of the partitive in Finnish.

Hopper and Thompson propose that transitivity is best understood as a composite of features pertaining to the whole clause. Laury (1982) shows that the partitive is mainly used in contexts that show features of low transitivity, and in Helasvuo (1996b) I have shown that the grammaticization of the partitive is strongly related to those features that Hopper and Thompson identify as characterizing low transitivity. Thus, with certain verbs of inherent low transitivity (the so-called “irresultative verbs” in Finnish linguistics), the partitive has become the only object marking possibility (ex. 29). The partitive has also been grammaticized to code object NPs under the scope of negation (ex. 30).

(29) Family

Liisa: *tuol fonetiika-n laitokse-l*
 there phonetics-GEN department-ADE
 At the phonetics department
puhu-u mone-t ihan murre-t.
 speak-3SG many-PL quite dialect-PTV
 many people speak some non-standard variety.

(30) Brothers

Pekka: *ei me saa-tu lapio-l mi-tä-ä,*
 NEG-3SG 1PL get-PCP spade-ADE nothing-PTV-PRON
 We couldn’t make any,
@@ mi-tä-ä reikä-ä sii-hen jäätynee-seen
 nothing-PTV-PRON hole-ILL 3SG-ILL frozen-ILL
 any hole in the frozen

lume-en
 snow-ILL
 snow with a spade.

In example (29), the object NP *murret* ‘dialect-PTV’ stands in the partitive, because the verb *puhuu* ‘speak’ is inherently low in transitivity. Here, no other case ending would be possible. In example (30), the object NP *mitää reikää* ‘any-PTV hole-PTV’ is in the partitive because it is under the scope of the negative verb *ei*.

Hopper and Thompson (1980) identify affectedness and individuation of the object as indicators of the degree of transitivity expressed by a clause. Thus, clauses in which the patient is highly individuated or strongly affected exhibit high transitivity. With their capacity to express open quantification, partitive NPs exhibit low transitivity. Partitive NPs are often mass nouns or they are in the plural, and thus, they are less individuated than are objects in the accusative or nominative case. Partitive objects may join the verb in a predicate phrase where the object is a separate NP which is not referential but predicating (ex. 31).

(31) Women

Eila: *ol-i aina tommos-i-s paiko-i-s kiva*
 be-PAST+3SG always that kind of-PL-INE place-PL-INE nice
teh-dä kauppa-a-ki
 do-INF business-PTV-CLT
 It was nice to shop in that kind of place.

In example (31) the object NP *kauppaa* ‘business-PTV’ forms a predicate together with the verb *teh-dä* ‘make, do’. The object NP *kauppaa* is not referring but predicating.

Crosslinguistically, the object role is clearly the slot for non-human referents, whereas the subject role typically accommodates human referents (see e.g. Givón 1983, Du Bois 1987b for Sakapulteko, Herring 1989 for Tamil, Ashby and Bentivoglio 1993 for French and Spanish, Nakayama and Ichihashi-Nakayama 1994 for Japanese, Kärkkäinen 1996 for English). This can be seen clearly in my Finnish data, where almost all transitive subjects refer to humans (over 90%), while less than half of the intransitive subjects and only 6% of the objects do. Given the fact that almost half of the partitive NPs function as objects in Finnish, we could hypothesize that the tendency for preferring non-human referents shows up in the discourse profile for partitive NPs. This is indeed the case, as can be seen in Table 9.

In Table 9, I have included all NPs irrespective of their syntactic function. We can see that both partitive and accusative NPs have very few human referents. Note that not only do partitive objects have a preference for non-human referents but partitive NPs functioning in other syntactic roles also prefer non-human referents. In contrast, over 40% of the nominative NPs refer to humans. Note also that half of the adessive NPs refer to humans (for further discussion, see Section 2.4.5).

Table 9. The distribution of human referents across cases.

Case of NP and semantic class	Human		Total
	N	%	N
Nominative	421	42.7	987
Accusative	3	4.3	69
Partitive	13	4.9	266
Genitive	15	31.2	48
Adessive	54	50.0	108
Other	40	12.6	318
Total	546	30.4	1796

In Helasvuo (1997b) I show that one of the main differences between the core and oblique cases in Finnish is the way in which they are used to keep track of referents in discourse (Durie 1994 reports on similar findings based on Acehnese conversational data). I claim that the core cases are used for introducing participants that will be tracked in the discourse, whereas only a few of the referents that are introduced using a locative NP in my data are ever mentioned again. Fewer than 10% of the referents that had been introduced with a locative NP were mentioned again. In other words, participants that were somehow central to the discourse were introduced in the core cases. In Helasvuo (1996b) I focused on the relationship between case marking and introduction of new referents. Table 10 summarizes some of the results by showing the case of the initial mention and the number of further mentions of the referents.

Table 10 shows that if a referent is introduced using a nominative NP, it is quite likely to be mentioned again (43%), whereas those referents that are introduced using the accusative case are less likely to be mentioned again (30%). The local cases rank lowest here: less than 10% of referents first mentioned in a local case were mentioned again. Interestingly enough, the partitive

is very close to the local cases here; less than 15% of the referents introduced with a partitive NP were further mentioned.

Table 10. The distribution of further mentions across cases with which the initial mention was made.

	Further mentioned		Total of new
	N	%	N
Nominative	112	43.4	258
Accusative	10	30.3	33
Partitive	19	14.6	130
Other	18	9.7	186
Total	159	26.2	607

It is important to note that this applies not only to partitive NPs functioning as objects but also to E-NPs (the so-called existential “subjects”). In Chapter 3, I show that although the Finnish existential constructions (constructions with a locative NP + *olla* ‘to be’ + E-NP) are used for introducing new referents, the referents of E-NPs are usually not mentioned again (see also Helasvuo 1996a).

As Table 9 shows, the partitive accommodates mainly non-human referents. From Table 10 we can see that partitive NPs are quite often used for introducing new referents, but these referents are not tracked, and they are not central to the discourse. We could claim that human referents in general are more central to our communication, and thus exhibit higher continuity in discourse (cf. articles in Givón 1983). Since both the accusative and the partitive have very low percentages for mentions of human referents (cf. Table 9), we would then expect them to show similar profiles with respect to the tracking function. However, referents that have been introduced with an accusative NP are much more likely to be mentioned again than the ones that are first mentioned with a partitive NP (cf. Table 10).¹⁴ A possible explanation for this is transitivity: in the course of its grammaticization process, the partitive has become strongly associated with low transitivity. Partitive NPs have referents that are less individuated and less affected than referents of accusative NPs. In

14. Note, however, that the absolute numbers for new mentions in accusative NPs are low, and it is possible that analysis of more data would give slightly different results.

clauses with partitive objects, the referents of the objects are not in focus but the processes are. It is only natural, then, that partitive NPs are not very likely to be tracked. In this respect, the partitive still shows characteristics that are more typical of oblique cases than the core cases.

2.4.4 Interim summary: The interplay between the grammatical cases in the coding of syntactic roles

In this section, I have shown that the distinction between the core and oblique cases is an idealization which is not very useful as a strict dichotomy, but rather as a fairly broad description. Instead of simply dividing the cases into two distinct sets, it is more accurate to say that some cases have been grammaticized to a greater extent than others.

The Finnish case system exemplifies the observed universal tendency for locative items to develop more grammatical meanings. The use of the nominative and the accusative is regulated mainly by grammatical factors, whereas the partitive carries more semantic meaning potential. As we have seen, this is understandable in light of the history of the case system, as the partitive was originally a local case.

The nominative occupies a central position in the case system, both in terms of structure and discourse use. Not only is the nominative unmarked phonologically and morphologically, but also in terms of several information flow features studied. The use of the nominative in discourse lends support to Du Bois' proposal that the unmarked category often has a special relation to preferred patterns of language use (see especially Du Bois 1987b). We found that most nominative NPs constitute non-new mentions. Moreover, they are usually used to track participants.

The other core cases, the accusative and the partitive, show distinctive discourse profiles. Both accommodate new mentions and most of their referents are non-human. However, in subsequent discourse, these referents that are introduced with accusative NPs are much more likely to be mentioned again than referents introduced with partitive NPs. Thus, in terms of their use in referent tracking, partitive NPs still have features that are more characteristic of oblique NPs than of NPs in the grammatical cases.

The case marking of the core roles, subject and object, has received a lot of attention in Finnish linguistics. I have shown that the seemingly complicated case marking patterns may be easier to explain if the marking of the personal pronouns is taken as the starting point. As we have seen, the personal pronouns

exhibit a straight-forward nominative-accusative marking system where the nominative is distinctively used to mark subjects. In contrast, the marking of other pronouns and full NPs is not as explicit as the nominative can be used to mark objects as well, if there is no subject in the clause.

2.4.5 The adessive and the other oblique cases

As we have seen in the previous sections, the distinction between core and oblique cases is not clear-cut. However, there are some discourse features that are more characteristic of NPs in the core cases than of oblique NPs. In this section, I will start out with a discussion of the discourse use of the oblique cases, with a special emphasis on the adessive case, and then I will discuss certain grammaticization processes in the oblique cases.

Oblique cases in discourse

In principle there are no strict rules governing the number of oblique NPs (adjuncts) per clause, whereas core arguments are much more restricted. Thus, one could hypothesize that oblique NPs would outnumber core NPs or at least occur equally frequently, but this turns out not to be so. In what follows, I will show that the flow of discourse is managed in the clause core, while obliques play a more marginal role for example in terms of referent manipulation.

As Table 2 showed, oblique NPs form a minority among case-marked NPs: less than 25% of NPs were marked with oblique cases. In comparison, Karlsson (1982:308) reports that in written discourse the percentage of obliques is slightly higher, 32.5%. In my spoken data, the adessive and the inessive were the most frequently used oblique cases (over 100 occurrences for each), whereas the elative and the illative were used a little less frequently (around 60 occurrences each). Some cases were hardly used at all in the data.

In discussions on the core-oblique distinction, the oblique cases are often described as being “circumstantial” (Halliday 1985) or “peripheral” (Foley and Van Valin 1984). According to Foley and Van Valin (1984:77), peripheral arguments express “the spatio-temporal setting of the event” or “secondary participants in the event, e.g. beneficiaries”. According to Halliday (1985:101, 137), “circumstantials” describe the “circumstances associated with the process” and are distinct from the participant roles (such as Actor and Goal) which express the participants of the process. These characterizations describe the functions of obliques from the viewpoint of a single sentence: obliques are somehow marginal in the sentence.

Interestingly enough, if we look at obliques from a discourse perspective, the same sense of marginality remains. In Helasvuo (1997b) I show that the main difference between core and oblique NPs is that core NPs are much more likely to be used to track participants than are obliques. Example (32) serves to illustrate this tendency. It comes from a series of stories about people's first visits to their girlfriend's or boyfriend's parents.

(32) Birthday

- 1 Leena: *ei mut mä oo-n kuullu kirjasto-ssa semmose-n*
 NEG+3SG but 1SG be-1SG hear-PCP library-INE such-ACC
 But in the library I have heard a story
- 2 *ku, joku oli menny, ... kanssa ensi-visiti-lle,*
 when somebody had gone also first-visit-ALL
 that (there was) somebody (who) had also gone for a first visit
- 3 *... poika-ystävä-nsä koti-in ja, ... sitte,*
 boy-friend-PX home-ILL and then
 to her boyfriend's home and then,
- 4 *... sit se o tietenki menny auttamaan tä-tä,*
 then 3SG had naturally gone help this-PTV
 then of course she had gone to help this,
- 5 *... (THROAT) toivottavasti tuleva-a anoppi-a*
 hopefully future-PTV mother.in.law-PTV
 hopefully future mother-in-law
- 6 *keittiö-ön ja,*
 kitchen-ILL and
 to the kitchen and,
- 7 *sit se oli huomannu,*
 then 3SG had noticed
 then she had noticed
- 8 *(H) ... to-ssa jää-kaapi-n ove-ssa lapu-n että,*
 that-INE ice-closet-GEN door-INE note-ACC that
 a note on the refrigerator door saying
- 9 *... huom, tämä ei ole Eija,*
 note this NEG+3SG be Eija
 "Note, this is not Eija,

- 10 ... *tämä on Tarja.*
 this is Tarja
 this is Tarja.”
- 11 Noora:@@@
- 12 Leena: ... *anopi-l* *oli muisti-lappu,*
 mother.in.law-ADE had memory-note
 The mother-in-law had a note (lit. memory note),
- 13 ... *bonjaa-k* *sä.*
 understand-2SG+Q 2SG
 do you get it?

In example (32) there are several new mentions, both in the core cases and in the oblique cases. I will first discuss the new mentions made in the core cases, and then I will move onto new mentions in the oblique cases. In line 2, the speaker introduces the main character of the story as *joku* ‘somebody’ in the intransitive subject role in the nominative case. The referent is mentioned again in line 7 with the 3rd person pronoun *se*. In lines 4 and 5, the speaker introduces another new referent in the object role with an NP in the partitive case *tätä toivottavasti tulevaa anoppia* ‘this hopefully to-be mother-in-law’. This referent is tracked in the discourse with a second mention in line 12 (*anopil* ‘mother-in-law-ADE’). In line 8, there are two NPs referring to referents that are new in the discourse, namely the inessive NP *tossa jääkaapin ovesa* ‘on the refrigerator door’ (for discussion, see below) and the accusative NP *lapun* ‘note-ACC’. The referent of the object NP, the note, is mentioned one more time in line 12, now with a compound noun *muistilappu* ‘note (lit. memory note)’ the first part of which (*muisti* ‘memory’) modifies the second part (*lappu* ‘note’). It is interesting to note that the second mention in line 12 is made using a form that is atypical of second mentions, since it is morphologically heavier than the first mention. In the story, the first mention is in the middle of the main storyline, whereas the second mention occurs in a sequence that reiterates the point of the story. Thus, the heavy morphological form of the second mention is iconic with its function as a reiteration. From the audiotape it is not possible to tell who the recipient is, but lines 12–13 are addressed to someone specifically (note the use of the 2nd person pronoun and person marking in line 13), and seem to be due to a lack of appropriate response on the part of the recipient at the end of the story.

In line 1, the NP *kirjastossa* ‘in the library’ is a first mention of the referent. It is made in the inessive case (ending *-ssa*). After this first mention the referent is never mentioned again. Similarly, the NP *poikaystävänsä kotiin* ‘to her

boyfriend's home' (line 3) introduces a new referent in the illative case, but the referent is never mentioned again. In line 8, there is a new mention made in the inessive case *tossa jääkaapin ovesa* 'on the refrigerator door (lit. in the door)'. The referent of this inessive NP, the refrigerator door, is not tracked any further.

The example shows that new mentions can be made with NPs in both core and oblique cases, but those referents that are introduced using oblique NPs are not tracked in the discourse; instead, they are dropped after their first mention.¹⁵ In contrast, those referents that have their first mention in a core case (*joku* 'someone' in line 2, *tätä toivottavasti tulevaa anoppia* 'this hopefully to-be mother-in-law-PTV' in lines 4 and 5 and *lapun* 'note-ACC' in line 7) are all tracked further. It may well be that it is characteristic of oblique NPs cross-linguistically that they are not usually used to track referents (Durie 1994 and Thompson (1997) offer supporting findings from other languages (Acehnese and English, respectively). Oblique NPs often provide settings for the events and states being described and it is quite natural that there is not such a great need to track anything in the settings since they are not usually manipulated, whereas the participants often are.

In Helasvuo (1997b), I discussed the distinction between the core and oblique cases and concluded that those referents that are being tracked, and thus are important in the discourse, are also important in the clause and belong to the clause core. However, this is a broad generalization that hides more delicate discourse profiles of the respective cases. As we saw in the discussion on the partitive, the so-called core cases do not behave uniformly with respect to the feature "tracking", and as we will see below, the oblique cases also show different profiles. As was discussed above, there is evidence of a universal tendency in languages for locative items to develop more grammatical meanings (see e.g. Heine et al. 1991, Harris & Campbell 1995). In this perspective, it is not surprising to find that the distinction between core and oblique cases is not clear-cut, but rather, represents a more general characterization according to

15. Tuomas Huumo (pc) notes that although the referents that are introduced with oblique NPs are not mentioned again, they may somehow implicate each other. In example (28), the boyfriend's home is not mentioned after its introduction, but instead we have a mention of the kitchen that is contained in the home. Furthermore, the kitchen is not mentioned anymore, but there is a mention of the fridge door that is located in the kitchen. In other words, there is a scope hierarchy in the locatives (see Huumo 1996c). I am grateful to Tuomas Huumo for discussing this example with me.

which some cases are more abstract and grammaticized whereas others have more semantic meaning potential.

Grammaticization processes in the oblique cases in Finnish

Grammaticization of the oblique cases has been the focus of much recent research in Finnish linguistics (cf. Leino 1993 on the elative, Onikki 1995, 1997 on the internal local cases, Inaba 1999 on the *n*-suffix). Huumo (1996a) has studied the development of the adessive case and noted several domain shifts in the course of its grammaticization process.

The adessive is claimed to have developed as a combination of the Uralic general local case (ending **-na*) and the locative derivational suffix *-la*. By Proto-Finnic, it had become part of the case system on a par with two other local cases, the ablative (separative meaning) and the allative (lative meaning), with which it formed a tripartite set, the so-called external local cases (cf. the so-called internal local cases; see e.g. L. Hakulinen 1979: 104, Huumo 1996a.) In addition to its local meaning the adessive has developed a possessive and an instrumental meaning (on the meaning change, see Huumo 1996a). Thus, in modern Finnish, the adessive is used in four domains, locative, temporal, possessive and instrumental.

In the locative and possessive usages, preverbal adessives have been grammaticized as parts of certain clausal schemas, namely, the existential and the possessive clause schemas (see Section 1.1). These schemas could be approximately represented as follows:

(33) Existential/possessive clause schema

$\text{NP}_{\text{LOC/POSS}} + \text{olla 'be'} + \text{NP}_{\text{NOM/PTV}}$

Note that in the existential construction, the preverbal locative need not be in the adessive, but can just as well be in another local case, most typically in the inessive. In a possessive clause, however, the preverbal item is in the adessive case (sometimes also ablative or allative, but not in the internal locative cases inessive, elative or illative). Consider examples (34a) and (34b).

(34) a. Alko

Pekka: *ai tontti-lla on oma kaivo.*

oh site-ADE be+3SG own well

Oh, you have a well of your own on the property (lit. on the property (there) is a well of (your) own).

b. Women

Saara: *hei-l on, mökki siin,*
3PL-ADE be+3SG cottage there
They have a cottage there.

In example (34a), the adessive NP *tontilla* names a locative space which is then characterized by “reporting its content”, the well (*kaivo*) (Huumo 1996a: 297). In (34b), the adessive NP *heil* ‘they-ADE’ expresses the possessor and the postverbal nominative NP *mökki* ‘cottage’ names the possessed.

Over half of all adessive NPs in the data appeared in possessive clauses that realize the schema in (33). The interpretation of an adessive NP as possessive rather than for example locative is very much dependent on the referential characteristics of the NP. Namely, all the possessive adessives in the data refer to human referents, most often to the speech act participants, whereas adessive NPs that are interpreted as locative often refer to referents for which the spatial dimensions are somehow prominent, such as areas (e.g. example (34a) *tontti* ‘(building) site’) or places. As we have already seen, NPs with human referents are treated in a very special way in discourse. One could hypothesize that the possessive adessives show discourse profiles that are different from those of other adessive NPs. This indeed is the case. With respect to the feature tracking, about half of adessive NPs are tracking, whereas for other oblique NPs, the percentage of tracking mentions is less than 20%. Also, adessive NPs are less likely to contain new mentions than other obliques. If we look at the adessive NPs more closely, we can see it is precisely the possessive adessives that make the adessive so different from other obliques in its discourse profile. It is also noteworthy that in the possessive adessives, personal pronouns figure significantly (almost 70% of possessive adessives are personal pronouns) whereas in other usages of the adessive and in other oblique cases personal pronouns are hardly used at all. This highlights the importance of personal pronouns in understanding how grammatical systems work in discourse (for further discussion, see Chapter 3).

From a grammaticization perspective the adessive is extremely interesting: the grammaticization of the adessive cannot be described without reference to the grammaticization of the existential/possessive construction as a whole. As Bybee et al. (1994: 11) note, linguistic elements are always grammaticized in constructions. The adessive has been grammaticized as part of the possessive/existential clause construction (see schema in 33), which is characterized by several crystallized patterns of language use that together make this clause type stand out as a type: the set of verbs that can be used is very limited (the verb *olla*

'be' and some other stative verbs), and so is the form of the two NPs (locative vs. nominative or partitive). Moreover, the word order is quite fixed with the locative NP in preverbal and the nominative/partitive NP (the E-NP, see Section 3.3.1) in postverbal position. Thus, there are several features that serve to distinguish this clause type.

In some respects, the possessive adessive resembles the subject: it has been noted that just like the subject, the possessive adessive can bind a reflexive in the postverbal NP (the possessed) (Hakanen 1972, Vilkkuna 1989). In terms of discourse properties, the possessive adessives resemble subject NPs in several respects: they have human referents, are often 1st or 2nd person pronouns, and they are usually tracking and do not contain new mentions.¹⁶ However, as I will show in Chapter 3, the grammaticization of the subject role follows a different path from the possessive adessive. The subject has been grammaticized as a syntactic role that is definable on the basis of syntactic features independent of construction type, whereas the grammaticization of the possessive adessive contains several features that are peculiar to the possessive construction type, such as the case marking of the two NPs, the strict word order and the severe restrictions on the verbs available in the constructions. In other words, in order to describe the grammaticization of the possessive adessive reference has to be made to the grammaticization of the possessive construction as a whole.

Summary of the oblique cases

In conclusion, we have encountered further evidence that the distinction between core and oblique cases is not clear-cut. It was shown, however, that the distinction is still useful in describing the functions of the cases in broad terms. We found that the oblique NPs differ from core NPs in that they do not usually track participants, but rather, they serve to orient speakers to the settings of the described events. As an example of grammaticization phenomena in the oblique cases, we looked at the use of the adessive in discourse and found that the adessive has been grammaticized as part of the existential/possessive clause schema. Those adessive NPs that are used in instantiations of this particular schema show a discourse profile that is quite different from other oblique NPs.

16. Vilkkuna (1996: 156–158) suggests that the possessive adessive could be analyzed as an "oblique subject". In my view, this is mistaken, since it lacks all the coding features that characterize subjects. Most notably, it does not trigger agreement, even if it is a 1st or 2nd person pronoun. As I will show in Section 2.5, 1st and 2nd person subjects always trigger agreement in the verb. For further discussion, see Chapter 3.

2.4.6 Summary of case marking

In our discussion of case marking in Finnish, we have seen that NPs marked with the core cases stand out as the primary means of tracking participants in discourse.

Several studies have shown that in spoken discourse there is rarely more than one full NP argument per clause. Therefore, pronominal coding patterns are central to the understanding of coding systems. However, since most previous research has been based either on single, isolated sentences or third person narrative, the role of personal pronouns in the emergence of coding patterns has gone underestimated. I have shown that the seemingly complicated case alternation of the core cases (nominative, accusative and partitive) is easier to explain if we take the case marking of the personal pronouns as a starting point. We have seen that the marking of the personal pronouns exhibits a classical nominative-accusative system, whereas the marking other pronouns and full NPs conforms to the nominative-accusative pattern to a lesser degree. The marking of plural full NPs follows a neutral alignment system. In the following sections, we will find further evidence for the central role of personal pronouns in the coding of syntactic functions.

2.5 Agreement

In the typological literature, agreement or cross-reference is often mentioned as one possible strategy for coding syntactic relations, on a par with case marking and word order (see e.g. Blake and Mallinson 1981, Croft 1991). In Finnish, agreement is employed in the coding of syntactic relations both on the clause level and within the noun phrase. In the clause, it codes the relationship between the subject and the verb; this could be called external agreement following Lehmann (1988:59), because the agreement refers to an NP outside the agreeing term (the verb). In the NP, agreement is employed in the coding of head — modifier relations, which Lehmann (1988) calls internal agreement. The NP internal agreement patterns will be discussed more closely in Chapter 4. In the present chapter, we will look at the agreement between the subject and the verb.

2.5.1 Sources of agreement markers

Pronominal anaphora is often mentioned as one of the most common diachronic sources of agreement (see e.g. Givón 1984, Lehmann 1988). Givón

(1984:360–362) assumes that there is a diachronic path that leads from independent pronouns through unstressed pronouns and cliticized pronouns to verbal cross-referencing morphology. In this section, we will look at the sources of agreement markers in Finnish. I will also discuss to what extent the original sources are still reflected in the present-day functions of agreement markers.

The development from independent pronouns to verbal agreement morphology has been attested in the Finno-Ugric languages, where the 1st and 2nd person verbal agreement markers indeed originate in personal pronouns. The original **mi-* and **ti-* pronominal stems (1st and 2nd person) are assumed to have been cliticized and later, to have become suffixes. During their grammaticization process, they have become eroded and subject to sound changes, eventually resulting *-n* for the 1st and *-t* for the 2nd person singular suffix. In the plural, the pronominal origins of the 1st and 2nd person suffixes are still clearly visible: the 1st person verbal suffix is *-mme* where the first *m-* stems from a present tense marker and the rest is the same as the 1st person plural pronoun *me*. The 2nd person suffix has gone through a similar development: the present tense marker is combined with the 2nd person plural pronoun *te*, resulting in *-tte*. However, the 3rd person markings are of a different origin: the 3rd person is either not marked at all (this is true of mono- and trisyllabic stems ending in a long vowel or a diphthong, such as *saa* ‘get+3sg’) or the marker stems from a present tense marker *-pa* (or in the lenited form *-va*), which is realized as a lengthening of the final vowel of the stem in the modern language, e.g. *istu-u* ‘sit-3sg’ (see below Table 11). In the 3rd person plural, there is a plural marker *-t* attached to the *-va*-ending. (See L. Hakulinen 1979.)¹⁷

It is important to note that the Finnish passive is also a personal form which is almost exclusively used to describe actions and states that involve an unspecified human agent or patient (for a discussion of the Finnish passive, see Shore 1988). In the passive form, there is a person marker (*-Vn*) on the verb in the morphotactic slot for personal suffixes (the so-called “4th person”; Tuomikoski 1971). This person marking is of the same origin as the 3rd person personal pronoun *hän* and the 3rd person personal marking in the reflexive inflection in certain dialects of Finnish (Koivisto 1995:61).

17. In some dialects of Finnish and in old written Finnish, the 3rd person plural is marked only with a *-t*-suffix in the past tense. In modern standard Finnish, the present tense marker *-va-* has been analogically extended to the past tense forms (L. Hakulinen 1979:247). In other words, *-va-* has been reanalyzed as belonging to the personal suffix instead of being analyzed as part of the tense marker.

Thus, the 1st and 2nd person markers originate in the respective pronoun, whereas the passive person marker stems from the 3rd person personal pronoun. In contrast, the person markings of the 3rd person are not pronominal in origin, but rather, the 3rd person either has no marking on the verb or is marked with a suffix which was originally a present tense marker.

Since the person markers in the 1st and 2nd person developed from the respective pronouns, there is double marking for person in these categories (cf. Du Bois 1997b on double positives). The verbal person marking indexes the referent of the subject, and it is possible to leave out the separate personal pronoun. However, while leaving out the personal pronoun in 1st and 2nd person is the preferred alternative in standard written Finnish, it is actually quite rare in conversation. It is interesting to note that in the 3rd person where the reference is far less clear (for reasons to be discussed in the following section), it is more common not to have an overt subject.

2.5.2 Agreement: the system and how it is used

In Finnish, verbs agree with the subject in person and number. Finnish distinguishes between 1st, 2nd and 3rd person and two grammatical numbers, singular and plural. The system is illustrated in Table 11; the inflection of the verb *katsoo* 'watch' is given as an example.

Table 11 shows that in the plural the colloquial system differs from the standard in several respects. In 1st person plural, the old personal suffix has been replaced by a passive form. This is a recent development, but it is quite pervasive in the colloquial language. In the 3rd person, number opposition is neutralized to a great extent, and the 3rd person singular form is used with plural subjects also. In my data, there were no instances with the 3rd person plural verbal suffix on verbs with 3rd person plural subjects; instead, the 3rd person singular form was used.¹⁸ Note, however, that although number opposition is neutralized in the 3rd person, it still shows person opposition.

The use of a passive form with the 1st person plural pronoun may seem odd at first sight. However, the combination of an indefinite form with 1st person plural reference has a counterpart in French, where the indefinite pronoun *on*

18. Based on a study of 3rd person plural agreement patterns in radio talk shows and call-in programs, Tiainen (1997) reports that plural agreement in the 3rd person is very rare in the speech of young reporters, but she also finds that it occurs more frequently in more formal programs and in programs of the national public radio.

is used for 1st person plural reference (see Laberge 1978, Ashby 1992). In French, the pronoun *on* is combined with the 3rd person verb form, and the interpretation of its reference as either indefinite or as a plural 1st person is construed in the context (Ashby 1992: 136). The Finnish pattern is different in that there is no agreement since the 1st person plural pronoun is combined with the passive form. As I will try to illustrate below (see ex. 35 and 36), the functioning of the passive and the plural 1st person are not all that different in creating referential relations in discourse.

The variety of forms and neutralization of oppositions found in the agreement patterns is interesting in view of the fact that agreement functions to index the subject. We could study the variation and development of agreement patterns to find evidence for the grammaticization of the subject role (for more discussion, see Chapter 3). In the 1st and 2nd person singular, the subject referent is easy to pin down. Interestingly enough, it is precisely in these forms that agreement works without exception to code the subject. In contrast, in the 3rd person and in the plural forms, there is more fuzziness, both in terms of agreement patterns and in terms of the referents of the subject NPs. For example, the set of referents included under *me* 'we' varies from context to context: of course, it always includes the speaker, but the rest of the referents are not necessarily clearly distinguished (see below ex. 36). Nunberg (1993: 7) notes, "It is not in doubt that *we* is a plural; what is less certain is that it is the plural of *I*." In the 3rd person, there is most variation in the possible subject referents: they range from human referents to inanimates, from concrete to abstract entities. It is noteworthy that in these forms agreement patterns are the least decisive. In Chapter 3, I will discuss the possible reasons for this patterning.

In principle, the 1st and 2nd person pronouns can be omitted when they function as subject, since the verbal suffix alone would be sufficient to convey

Table 11. Agreement system in Finnish. Verbal agreement morphology.

Person	Singular	Plural	
		standard system	colloquial system
1st person	- <i>n</i> (<i>katson</i>)	- <i>mme</i> (<i>katsomme</i>)	PASS: - <i>tVVn</i> (<i>katsotaan</i>)
2nd person	- <i>t</i> (<i>katsot</i>)	- <i>tte</i> (<i>katsotte</i>)	- <i>tte</i> (<i>katsotte</i>)
3rd person	- <i>V</i> (<i>katsoo</i>)	- <i>vAt</i> (<i>katsovat</i>)	- <i>V</i> (<i>katsoo</i>)
passive	- <i>tVVn</i> (<i>katsotaan</i>)		

the information about person (e.g. *istun* 'sit-1SG' can be used instead of *minä istun* 'I sit-1SG' to convey the meaning 'I sit/am sitting'). In the standard language, this is the preferred alternative, whereas in the colloquial language, it is rare to have the 1st or 2nd person pronoun omitted; in my data, there were only a couple of instances. In clauses with 3rd person verb form, if there is no pronoun, the zero may be interpreted as anaphorically bound to some previous mention or as indexically open so that the reference can be co-constructed by the conversation participants (see Laitinen 1995, 1996).

I will try to illustrate the construal of personal reference with the following example, which comes from the Alko conversation where the men are talking about how one can save money by quitting smoking. Aarre has just stopped smoking, and now considers the pros and cons.

(35) Alko

1 Aarre: *mu-ll on suuremma-t ^meno-t-ki tullee-vat*
 mv1SG-ADE is larger-PL expense-PL-CLT come-3PL
 I have more expenses/the expenses become more

2 [*helevetti. et,*]
 damn.it that
 damn it. So,

3 Pekka: [*joo, näin on.*]
 yeah so is
 Yeah, you are right.

4 Aarre: *^voi-ta-s-ha me se,*
 can-PSS-COND-CLT 1PL it
 Sure we could (do) it,

5 *ei-hä si-tä voi sano-o kukkaa,*
 NEG+3SG-CLT it-PTV can say-INF anybody
 nobody can deny that

6 *että yhtä hyvi-hä si-tä vois pärjä-tä että,*
 that a.like well-CLT it-PTV could manage-INF that
 one could manage just as well

7 *vaikka nyt ei,*
 although now NEG
 even if now

8 Pekka: *joo=.*
 Yeah

- 9 Aarre: *minä o-is* *^lopetta-nuk-ka si-tä,*
 1SG be-COND quit-PCP-CLT it-PTV
 I wouldn't have quit it (i.e. smoking).
- 10 ... *mutta aina* *^paremmi pärjä-tä-än* *sillai et <X X>.*
 but always better manage-PSS-PERS so that
 but (we) will manage even better this way...

In example (35) in line 1, we have a double bind structure (Franck 1985) that starts out as a possessive construction (*mull on suuremmat menot* 'I have larger expenses') but at the same time, the NP referring to the possessed entity, *suuremmat menot* 'larger expenses', acts as the subject of the verb *tulleevat* 'become', which is one of the rare examples in the data of 3rd person plural agreement in number. In line 4, Aarre starts with the construction *voitasha me* 'we could', in which a verb with a passive form combines with a 1st person plural subject (see below). In line 5 Aarre shifts the focus to a more general level and uses the indefinite pronoun *kukkaa* 'anybody' as a subject. In line 6 we have a 3rd person verb form *vois* 'could' with a zero subject. The zero is indexically open: of course the speaker is primarily talking about himself and his family, but he formulates his statement so that it allows the co-participants to identify themselves with the indexically open reference (see Laitinen 1995, 1996). After this point, Aarre first shifts the deictic center to himself in line 9 by using a 1st person singular pronoun, and then back to a more open reference in line 10 by using the passive *pärjätään* 'manage-PSS'. Again it is quite obvious that Aarre is talking about his own family, but he uses a form that leaves the reference open.

Example (36) further illustrates the agreement patterns. It comes from a telephone conversation between two brothers, Pekka and Simo. Pekka has called his brother to tell about his fishing trip to Lapland. They were supposed to overnight in a tent, but it has turned out that two Norwegians had invited them to share a cabin with them. (Lines 6–8 and 11 do not have a morphological gloss and will not be discussed in detail. They have been included only to show how reference is created and maintained during the discourse.)

- (36) Brothers
- 1 Pekka: *me ol-t-i-i* *vaan siel*
 1PL be-PSS-PAST-PERS just there
 We were just there (sitting)
- öljy-kamina-n lämmö-ssä sitte.*
 oil-heater-GEN warmth-INE then
 in the warmth of an oil heater.

- 2 Simo: ... *^ai jaa*,
oh yeah
Oh I see,
- 3 *te oo-tte ollu ihan herro-i-ks sit siäl.*
2PL be-2PL been quite sir-PL-TRA then there
you have had it quite comfortable there. (Lit. "You have
been quite like sirs there.")
- 4 Pekka: ... *^joo*.
yeah
- 5 Simo: *ol-i-ks ne ^norjalais-i-a.*
be+3SG-PAST-Q they Norwegian-PL-PTV
Were they Norwegians?
- 6 Pekka: ... *no se oli niinku, uh=, ... ^Ruotsi-n norjalainen tää toinen*
ukko ja,
Well he was like, uhm ... a Sweden Norwegian this other
guy and
- 7 *toinen oli ihan sellane, ... (H) ... ^aito norjalainen.*
the other one was like, ... (H) ... a real Norwegian.
- 8 Simo: *oliks ne kaks siäl ^kämpillä.*
Were the two (staying) there in the cabin?
- 9 Pekka: *^joo. ne ol-i siin ^toise-l puole-lla.*
yeah 3PL be-PAST+3SG there other-ADE side-ADE
Yeah. They were there on the other side.
- 10 *ja me ol-t-i-in siin ^toise-l puole-l.*
and 1PL be-PSS-PAST-PERS there other-ADE side-ADE
And we were there on the other side.
- 11 Simo: *^ai jaha. ... no mites,*
Oh, okay. Well how was it,
- 12 Simo: *pääsi-tte-ks te kalastele-maa.*
get-2PL-Q 2PL fish-INF
Were you able to go fishing?
- 13 Pekka: *^joo.*
yeah
- 14 Simo: ... *^tul-i-ko rautu-a.*
come-3SG+PAST-Q trout-PTV
Did you catch any trout?

- 15 Pekka: *no ku teh-t-i-i* *^ensimmäine reikä,*
 well when make-PSS-PAST-PERS first hole
 Well, when we made the first hole (in the ice),
- 16 *ja mä tiput-i-n* *se-n lipa-n sinne,*
 and 1SG drop-PAST-1SG it-ACC lure-ACC there
 and I dropped the (fishing) lure in there,

In example (36) line 1, Pekka refers to himself and his fishing partners with the pronoun *me* ‘we’. One of the partners has been mentioned by name during the conversation, but it seems that they were more than just the two. The brother may or may not know who exactly the ‘we’ are. The verb stands in the passive form. In line 3, we have a 2nd person plural pronoun *te* functioning as a subject, and the verb *ootte ollu* ‘have been’ agrees with it in person and number. The reference of the subject pronoun *te* is open for interpretation: it can be understood as being coreferential with *me* in line 1. In line 5, the subject of the copular clause *ne* ‘they’ triggers person agreement in the copular verb *oli* ‘was’, but there is no number agreement. Here the subject pronoun *ne* is coreferential with a previous mention not included in the example. The same referent is mentioned again in line 9, again with the pronoun *ne*. In line 10, the first person plural pronoun *me* ‘we’ is used as a subject to refer to Pekka and his fishing partners.

After Pekka has told Simo about the two men and the sleeping arrangement that they had had in the cabin (lines 6–10), Simo asks whether they were able to go fishing (lines 11–12) despite the cold and stormy weather. Pekka gives a positive response (line 13), and after a small pause, Simo continues to ask whether they caught any fish (line 14). In his first question (line 12) Simo uses the 2nd person plural. This is an interesting choice referentially: it could refer to Pekka and his fishing partners or it could include the Norwegian men that were sharing the cabin. Interestingly enough, in his answer in lines 15 and 16, Pekka does not use the first person pronoun *me* that he had been using prior to this (lines 1 and 10), but only the verb in the passive form. The clause could be analyzed as containing zero anaphora coreferential with the 2nd person plural pronoun *te* in line 12, but the exact set of referents is not specified because *te* is referentially ambiguous. Or, alternatively, the clause in line 15 could be analyzed as containing a passive with an unspecified agent. It seems, however, that the two analyses are not very different from each other: in both cases, the reference is contextual and has to be inferred.

From examples (35) and (36) we can see that there is certain parallelism in the use of the 3rd person singular verb form with a zero subject and the patterning of the 1st person plural. In both the 3rd person singular and in the 1st person plural, if there is no pronoun and we only have the verb form (verb + 3rd person singular ending or verb + passive ending), the clause may be interpreted in (at least) three ways:

- i. as containing zero anaphora (see ex. 36),
- ii. as containing an indexically open zero (see ex. 35), or
- iii. if the clause is verb-initial, as a suggestion (optative, e.g. *syödään* 'Let's eat').

These alternatives are available in both the 3rd person singular and in the 1st person plural, if there is no pronoun. However, the third alternative interpretation (i.e., an optative) is somewhat archaic and marginal in the case of the 3rd person singular (*Tulee nyt vaan* 'come-3sg now just', a polite way of saying 'Come on'). Of course there are also differences between the two forms: for example, in connection with the passive, the indexically open zero (item (ii) above) is (almost) always human and agentive, whereas with the 3rd person singular verb form, the zero is often not agentive, but rather, an experiencer (see Laitinen 1995b). Also, in connection with the 3rd person singular verb form, if the zero is anaphorically bound to some previous mention, the anaphor can refer to a wide variety of entities, whereas with the passive form, the anaphoric zero will always have 1st person plural reference.

Thus, omission of the pronoun has very different effects in connection with the different personal forms of the verb. The 1st and 2nd person singular and the 2nd person plural forms are still personal forms even if the pronoun has been omitted, whereas the 3rd person and the 1st person plural may turn into forms with which the personal reference is not stated, but rather, it has to be inferred.

2.5.3 Cross-referencing or indexing

Lehmann (1988) emphasizes the referential nature of agreement morphology; thus, we may speak about cross-referencing morphology. According to Lehmann (1988:55), agreement serves to help to identify or reidentify referents. I will discuss this claim in this section and show that it is somewhat problematic, especially in connection with the personal verb forms with zero subjects.

In the 1st and 2nd person singular the agreement morphology works in a quite straight-forward way to cross-reference the 1st or 2nd person subject, be

it zero or overt. In both the singular and the plural, 1st and 2nd person index the speech act participants (speaker and addressee). In the plural, however, the reference may not be as clear-cut as is often the case with the singular forms; rather, it has to be inferred from the context which referents are included. Example (36) (esp. line 12) illustrates this (repeated here as 37).

(37) Brothers

12 Simo: *pääsi-tte-ks te kalastele-maa.*
 get-2PL-Q 2PL fish-INF
 Were you able to go fishing?

As was noted above, the use of the 2nd person plural pronoun *te* as the subject pronoun allows for different interpretations as to which referents are intended, whether it includes just the brother and his fishing partners or also the Norwegian men they met on their way. In the sequence chosen for example (36) we saw that this inexplicitness can be used as a resource in conversation.

We could say that in the 1st and 2nd person singular and the 2nd person plural, the agreement morphology serves to identify or reidentify the same referent(s) as the respective pronoun does, whether or not the pronoun is there. It is interesting to note that in these forms, the verbal person marking originates in the respective pronoun, and in the case of the 2nd person plural, it is almost identical with it (the verbal person suffix is *-tte* and the pronoun is *te*, see ex. 37). In the remaining forms (1st person plural and 3rd person singular and plural), the reference of the person marking is open to a wide range of interpretations if there is no pronoun. At the same time, in these forms the agreement markers do not originate in the respective pronoun, but rather, the markers are either not pronominal at all (3rd person) or they have their origins in a different person (1st person plural and passive).

In the 1st person plural, the question of reference is even more complicated. We can say that the person marking (the passive) is referential, but the reference is highly context-dependent. If there is a 1st person plural personal pronoun, the person marking on the verb is co-referential with it. But if there is no pronoun, there are several different ways to interpret the reference: it may be anaphorically bound to some previous mention, or it may be indexically open so that its reference has to be inferred from the context or it may function as an optative that includes the speaker. However, these alternatives are not choices that would all be available to the participants at the moment of speaking but rather, the interpretation is constrained by the context, and we could say that it is co-constructed by the participants in the context.

In the 3rd person singular and plural, the set of possible subjects and their referents is open-ended, and there are no semantic or discourse features that they would share. It is not surprising then that if there is no pronominal or full NP subject, the reference of the verbal person marking may be interpreted in a variety of ways.

I would like to further emphasize that this is true not only of these forms but of personal reference in general. Reference is contextually construed not only in the 3rd person and the plural forms, but also in the 1st person singular and 2nd person plural: there is no *I* except in the discourse context, and the interpretation of *you* is dependent on the current participant framework (in the sense of Goffman 1981). Furthermore, it is important to note that even when referents are being tracked with the help of anaphoric elements such as pronouns, the qualities of the referent do not remain stable while it is being tracked. Rather, the participants are constantly updating their understanding of the intended referent, and it is not just the referential expressions but also the predications that are being made about the referents that serve to add to that understanding. As we saw in the analysis of example (36), the reference of the 2nd person plural pronoun *te* does not necessarily remain the same although the referent is being tracked (cf. the use of *me* ‘we’ and *te* ‘you (PL)’ in lines 1 and 3 in comparison with *te* in line 12). Example (36) further illustrated that referential ambiguity and openness is well tolerated in discourse and does not cause any signs of trouble on the part of the co-participants (no hesitations, repairs etc.). To the contrary, it can be exploited as an interactional resource.

Thus, Lehmann’s claim that agreement marking “helps identify or reidentify referents” (Lehmann 1988) seems to be too simplistic to account for the functions of agreement in discourse. I would like to propose that verbal agreement marking is indexical, just like personal pronouns are indexical (see Nunberg 1993). In the 1st and 2nd person singular, the agreement marker serves to index the speaker and the recipient, respectively, whether or not the pronoun is there. In the 2nd person plural, the agreement marker indexes the recipients, but it has to be inferred from the context which referents are included. In the 3rd person and 1st person plural, verbal agreement marking serves to index a referential slot, and if there is no NP which it could provide an index for, the index remains open and has to be constructed in the context.

2.5.4 Summary: Agreement in the coding of grammatical relations

Agreement is related to case marking in several ways. First, only nominative NPs can trigger agreement in the verb. As we will see in Chapter 3, partitive NPs are sometimes used very much like subjects, but nevertheless, they lack the capacity to trigger agreement. Secondly, when the 1st person plural pronoun is used with a verb in the passive form, the object is marked as if the clause was a passive clause and thus, it can stand in the nominative, e.g. *me ostetaan auto* ‘we (will) buy a car’ (cf. with a real passive clause *auto ostetaan käteisellä* ‘a car is bought with cash’). In both clauses, the object *auto* ‘car’ is in the nominative, although in the first clause there is a subject NP in the clause as well (cf. with a clause with a subject NP *Mari ostaa auton* ‘Mari will buy a car’ where the object NP has accusative case). Here case marking works in a reverse way: the fact that the personal pronoun *me* is in the nominative shows that it cannot function as the object, because as we saw in Section 2.4.2, personal pronouns are always marked with the *t*-accusative when they function as objects.

To sum up, agreement works to index the subject. In the 1st and 2nd person singular and in the 2nd person plural, the coding is explicit and shows the person and number of the subject. In the 3rd person, however, the coding is far less explicit, since the number opposition and the opposition between human and non-human are neutralized. In the 1st person plural, the colloquial language shows an interesting development where the passive which is a personal form referring to a human agent has come to be used with the 1st person pronominal subject, i.e. a speech act pronoun. This development may seem odd at first, but if we consider the referential characteristics of the first person plural pronoun, we can note that its reference can hardly ever be determined explicitly, but rather, it almost always has to be inferred from the context. In view of this, it is not so surprising to find that it co-occurs with the personal passive.

2.6 Word order

Finnish word order has often been characterized as being relatively free with “very few genuinely grammatical constraints” (Vilkuna 1989: 9–10). If we look at word order from the perspective of single sentences or isolated fragments of texts, we can note that syntactic permutations rarely yield ungrammatical

orderings.¹⁹ However, if we take language use as a starting point, a very different pattern emerges. As we will see, the tendency for the subject to precede the verb is very robust, especially in spoken discourse. I will further demonstrate that the category of person, or, more broadly, the Silversteinian “inherent lexical content” of the subject noun phrase is extremely important to the understanding of word order patternings. We will also look at the distribution of the different types of subject NPs across clause types.

2.6.1 Word order and the preverbal position

In her study of word order in Finnish, Vilkuna (1989) has highlighted the importance of the preverbal position in the clause, both in terms of syntactic organization and in terms of information structure. In this section, I will discuss Vilkuna’s proposal as well as other studies on word order in Finnish and compare them to the patterns found in my data.

According to Vilkuna (1989), the preverbal position is the place for the topic or theme of the clause (Vilkuna’s “T”-element). Syntactically it is normally occupied by the subject, but in certain clause types some other element that regularly occurs before the verb, *inter alia* the locative NP in an existential clause (see Chapter 3). Of course, Vilkuna is not the only one who has emphasized the centrality of the preverbal position both in syntactic organization and in management of information flow (see e.g. Halliday 1985 on the conflation of subject and theme).

Several corpus-based studies have shown that there is a clear tendency in Finnish for the subject to precede the verb and for the object to occur after the verb. In Hakulinen et al.’s corpus of written Finnish, the subject preceded the verb in 61% of the clauses (Hakulinen et al. 1980: 145), and the same tendency can be seen in Huumo’s data from literary fiction where the percentage lies around 77% (Huumo 1994). In my spoken data, over 90% of the subjects

19. In her in-depth analysis of Finnish word order, Vilkuna (1989) uses naturally occurring data to illustrate her claims. However, her study is not corpus-based in the sense that she would attempt to describe the patternings found in a specific corpus. Furthermore, she has specifically sought example sentences with unusual orderings, and in her work, these exceptional cases receive far more attention than the prevalent patterns (Vilkuna 1989: 9–10). Thus, the findings presented here are in no way contradictory with hers, since her research goals are completely different from mine.

preceded the verb.²⁰ This tendency is even more robust if we look at pronominal subjects only. If we further separate pronouns referring to humans, we come fairly close to a fixed Subject–Verb pattern.

The following example (ex. 38) illustrates the word order patterns found in the data. In the example, Linnea is telling about her summer cottage and the problems with the sea water there: it grows a weed that Linnea and her family are constantly trying to get rid of.

(38) Women

- 1 Linnea: *kato muute ei si-tä saa*
see otherwise NEG+3SG it-PTV get+INF
See, otherwise one cannot get it (i.e. the weed)
- 2 *^mi-llä [-än siält sitte ylös]*
nothing-ADE-CLT from.there then up
up in any way.
- 3 Saara: [*no ^tule-e-k se siält ylös*] *sitte kaikki*.
so come-3SG-Q it from.there up then all
So does it all come up from there (i.e. from the sea)?
- 4 Linnea: *se tule-e ^juuri-ne-s sit kato*,
it come-3SG root-COM-3PLPX then see
It comes loose from the root then you see,
- 5 [*ku se saa*,]
when it get+3SG
when it gets,
- 6 Saara: [*tule-e vai*.]
come-3SG or
Really? (more literally: Oh it comes.)
- 7 Linnea: *^joo. ... (0.7) kyl se vähä autta-a mut*,
yeah yes it little help-3SG but
Yeah. It does help a little but,

20. The numbers are not fully comparable because of differences in the definition of subject. In my work, subject is defined as the core argument of transitive, intransitive and copular verb that stands in the nominative case and triggers agreement in the predicate. Hakulinen et al. (1980) and Huomo (1994) apply a broader definition that includes e.g. the core NP in an existential construction (the E-NP, see Chapter 3).

- 8 (H) *se o, ei me tul-tu* [^]*ajatel-leeks si-tä*
 it is NEG 1PL come-PSS+PCP think-INF it-PTV
 it is, we didn't come to think about it
- 9 *kato ku me ruet-t-i-i* [^]*ruappaa-maa et,*
 see when we start-PSS-PAST-PERS dredge-INF that
 you see when we started to dredge it that

In example (38), we have two clauses with zero subjects: in the clause in lines 1 and 2, the zero subject is indexically open (see above Section 2.5) but in the clause in line 6, it is anaphorically bound to the subject pronoun *se* 'it' in Linnea's clause (line 4). In all other clauses the subjects are pronominal (the pronoun *se* 'it' in lines 3, 4, 5, and 7, and in the short incomplete clause in the beginning of line 8 and the pronoun *me* 'we' in lines 8 and 9). Thus, in the example, the pronominal subject usually precedes the verb, with the exception of the clauses in lines 3 and 8.

In line 3 we have a question which is formed with the question clitic *-k(O)* attached to the word being queried, in this case *tulee* 'comes'. The word carrying the *-kO*-clitic almost always appears in the first position in the clause, and this of course alters the word order patterns of question clauses. In line 8, we have the negative verb *ei* preceding the subject *me* 'we'. In Finnish, the negation element has verbal features, as it inflects for person. However, it is not a full verb — it does not inflect for tense for example — and in some usages it is closer to being a particle (for discussion, see Lindén 1963). In many dialects, even the person inflection is lacking (Savijärvi 1977). The negation element has particle-like features with regard to word order as well: it seems to prefer the beginning of clauses or intonation units just like particles do.

Crosslinguistically, negative clauses and questions have been shown to behave differently with regard to word order from other types of clauses (see Dahl 1979, J. Payne 1985). In the tables presented in the following section I have excluded questions and clauses with negation. They deserve a separate study which unfortunately lies outside the scope of the present study.

In sum, Vilkuna (1989) has emphasized the importance of the preverbal position both in terms of the syntactic and informational structuring of the clause. When we look at actual discourse, it seems that the preverbal position is even more uniform than one would expect when looking at isolated example sentences.

2.6.2 Word order and type of subject

In this section, we will look at the mutual ordering of subject and verb in the data. I will show that the ordering of subject and verb is quite fixed (Subject–Verb) in clauses containing pronominal subjects, whereas full NP subjects are more flexible with regard to word order. These differences will be illustrated with Tables 12–14, to be followed by a discussion of the findings.²¹

First, we will look at 3rd person pronominal subjects and their distribution with respect to the verb. This is illustrated in Table 12, which shows that 3rd person pronominal subjects are mostly in preverbal position. There does not seem to be much difference between human and non-human referents in the percentages — both have percentages around 95% for subjects preceding the verb.

Table 13 below looks at word order in clauses with 1st and 2nd person subjects. It shows that over 96% of 1st and 2nd person pronouns also precede the verb when they function as subjects. If we look at the absolute numbers, we can note that the majority of 1st and 2nd person subjects are subjects of transitive clauses (66%) and only a few function as subjects of copular clauses (12%).

In sum, 1st, 2nd and 3rd person pronouns work similarly for word order as subjects almost always precede the verb. However, 1st and 2nd person subject pronouns show a different distribution across clause types from that of 3rd person pronominal subjects: 1st and 2nd person subjects are usually subjects of

Table 12. Word order in clauses with 3rd person pronominal subjects.

		Before verb		Total
3rd person pronouns		N	%	N
Human	St	42	91.3	46
	Sc	27	100	27
	Si	21	100	21
	Total	90	95.7	94
Non-human	St	11	91.7	12
	Sc	111	93.3	119
	Si	52	96.3	54
	Total	174	94.1	185

21. Note that questions and clauses with negation or with zero subjects are excluded from Tables 12–14. In the tables, St = subject of a transitive clause, Sc = subject of a copular clause, Si = subject of an intransitive clause.

Table 13. Word order in clauses with 1st or 2nd person subjects.

1st and 2nd person	Before verb		Total
	N	%	N
St	99	96.1	103
Sc	19	100	19
Si	32	94.1	34
Totl	150	96.2	156

transitive clauses, whereas for 3rd person pronominal subjects referring to non-human referents, the majority appear in copular clauses. Clauses with 3rd person pronominal subjects referring to human referents show a preference for transitivity as almost half of the clauses are transitive (47%), but the preference is not as clear as in the case of clauses with 1st or 2nd person subjects (66%).

In contrast, when we look at clauses with full NP subjects, we no longer find a strict Subject–Verb ordering, but instead, there is more flexibility. Consider Table 14.

As Table 14 shows, in clauses with full NP subjects referring to humans, 84% of the subjects precede the verb. In contrast, for full NP subjects that refer to non-humans, the average is only 63%. Note also the low absolute numbers for full NPs in comparison to the numbers for pronominal subjects. If we compare Tables 12 and 13 to Table 14, we can see that pronominal subjects are much more frequent than full NPs (79% of subjects are pronominal).

Table 14. Word order in clauses with full NP subjects.

Full NPs		Before verb		Total
		N	%	N
Human	St	28	87.5	32
	Sc	9	75.0	12
	Si	10	83.3	12
Total		47	83.9	56
Non-human	St	3	75.0	4
	Sc	15	75.0	20
	Si	21	55.3	38
Total		39	62.9	62

Tables 13 and 14 show that the ordering of 1st and 2nd person subjects with respect to the verb is much more constrained than that of full NP subjects. But we can also see that 1st and 2nd person subjects are most often subjects of transitive clauses, whereas in the other subject types there is no such clear preference.

The tables show clearly that the majority of subjects are pronouns rather than full NPs, which is in line with the findings of several discourse studies (see e.g. Lambrecht 1987, Chafe 1994, and also Du Bois 1987a). If we compare Tables 12 and 14, we can see that non-human subjects are most often subjects of copular clauses (see Chapter 3 for further discussion).

In conclusion, I have shown that the ordering of subject and verb is quite fixed in clauses with 1st and 2nd person subjects, whereas in the 3rd person, there is more variation. In clauses with 3rd person pronominal subjects, subjects precede the verb in about 95% of the clauses. In contrast, clauses with full NP subjects show a pattern which is more flexible, especially if the NP is referring to a non-human referent.

2.6.3 Summary on word order

In our discussion of word order, we have encountered further evidence for the tightness of the bond between the subject and the verb, especially in the case of 1st and 2nd person subjects. In Chapter 5 I will discuss the interplay between syntax and intonation, and I will show that subjects are almost always in the same intonation unit as the predicate verb. This is of course natural in the light of our findings on the fixed word order between the subject and the verb.

As I have shown in the previous sections, case marking and agreement work to single out 1st and 2nd person subjects with strict codings, whereas 3rd person subjects have less explicit codings. In this respect, word order patterns in a slightly different way: the 1st and 2nd person work similarly with the 3rd person pronominal subjects in showing fixed word order, whereas full NP subjects are much less constrained in their word order.

2.7 Summary: The clause revisited

Our discussion of case marking, agreement and word order has shown that the three strategies work together to code syntactic functions. They interact with each other in several ways: for example, if case marking does not distinguish

between the subject and the object, then word order may help to interpret the syntactic relations. Also, case marking works together with agreement so that only nominative NPs can trigger agreement. And finally, word order interacts with agreement so that postverbal subjects do not always trigger agreement even in the standard written language. It is important to note, however, that it is quite rarely that one finds cases in actual discourse where case marking would not give hints for the interpretation of syntactic roles.

Our discussion of word order showed that most subjects are preverbal, and often 1st or 2nd person pronouns, but if a subject NP is postverbal, it is most likely in the 3rd person. This has interesting implications for agreement patterns also: as we have seen, 1st and 2nd person subjects trigger agreement in the verb both in number and in person, whereas in the 3rd person, agreement is less decisive as a subject-coding strategy, as number opposition is neutralized to a great extent. Thus, postverbal subjects, being most often 3rd person pronouns or full NPs, lack explicit coding for the subject role not only in terms of agreement but also in terms of word order. In sum, we could say that the three coding strategies work together supporting each other.

Our discussion has highlighted the importance of the category of person for the understanding of how the coding systems work. In the case system the marking of personal pronouns provides a new perspective on case marking of syntactic relations in the clause core, especially the marking of the object. Regarding agreement, the agreement patterns are most explicit in the 1st and 2nd person but in the 3rd person there is more indeterminacy. And finally, word order in clauses where the subject is a personal pronoun shows a fixed Subject–Verb ordering, whereas in clauses with a full NP subject, there is more flexibility.

In previous research, the importance of personal pronouns has been largely underestimated. Several studies have acknowledged that personal pronouns show coding patterns that differ from those of other pronouns and full NPs, but this finding has often led to the exclusion of personal pronouns from the description of coding systems, especially in studies of word order. Even studies based on actual discourse may be somewhat biased in this respect if the data come only from third person narrative. Instead of being marginalized, personal pronouns should be taken as a central force in the structuring of argument relations in discourse.

Our discussion of the case system has shown that the core cases do important discourse work in the marking of NPs that serve to track participants in discourse. The core cases interact with each other in many interesting ways: for example,

if there is no nominative subject in the clause, the object can stand in the nominative. In the oblique cases, however, there is no such interaction. Agreement patterns also serve to distinguish the core roles from each other. Finally, in our discussion of word order, we looked at the core roles and their mutual ordering. Thus, all three coding strategies discussed here point to the clause core as the locus of the structuring of syntactic relations. Furthermore, the clause core was shown to play a central role in managing information, *inter alia* in delivering new information or tracking participants in discourse.

In conclusion, we have found a lot of evidence for the clause as an emergent grammatical category in Finnish. This is interesting in cross-linguistic perspective. As I mentioned, Tao (1996) has shown that the clause does not emerge so clearly as a level of syntactic organization in Mandarin discourse; instead, syntax is centered around the phrase. On the other hand, Nakayama (1997) has shown that in *Nuu-chah-nulth* (Nootka), syntactic structuring emerges on the word level rather than, say, in argument relations between the predicate and its arguments realized as clauses as in Finnish. Thus, languages respond in different ways to the various discourse needs that speakers are faced with, such as the need to express events or track participants of an event. On the basis of careful discourse studies on individual languages, we can start to build a cross-linguistic understanding on what clauses are and what kind of work they do in different languages.

CHAPTER 3

Grammaticicization of the subject role

Our discussion of the different coding strategies for distinguishing grammatical roles has been based on the assumption that only NPs which are in the nominative and which trigger agreement should be considered subjects in Finnish. In this section, I discuss the subject role from the point of view of grammaticization. First, I will address the question of whether the first arguments of transitive, intransitive and copular clauses really represent a unified subject role in Finnish. Then I will discuss the grammaticization of the subject role in terms of information flow features that characterize subject NPs, and finally, I will contrast it with two types of NPs which have traditionally been considered to be subjects in Finnish linguistics, namely the so-called existential “subjects” and “subjects” with open quantification.

3.1 Is there a unified subject role in Finnish?

In the discussion so far, I have maintained that the subject in Finnish should be defined in simple morphosyntactic terms. The different codings (nominative case, agreement) are the same for subjects of transitive, intransitive and copular clauses. However, as I have already hinted in the previous sections, there are differences in the distribution of different types of subject NPs across clause types. In this section, I would like to discuss whether we can still talk about a unified subject role in Finnish. I will also show how and why my definition of subject differs from the received view on subject in Finnish linguistics.

3.1.1 Subject and clause type

Our discussion of the use of the different coding strategies in discourse has shown that the subject is most explicitly coded when it is a personal pronoun, whereas full NPs receive less explicit codings. The coding principles apply alike to subjects of transitive, intransitive and copular clauses. However, as we saw in the section on word order, the different subject types (personal pronouns vs.

other pronouns, pronominal vs. full NPs, NPs referring to humans vs. non-humans) are not evenly distributed across clause types (transitive, intransitive and copular clauses).

As was pointed out in the section on word order, the majority of transitive subjects are personal pronouns, whereas full NPs referring to non-humans rarely function as transitive subjects (see Section 2.6.2, Tables 12–14). This means that in transitive subjects we find the most explicit codings. In what follows, we will discuss the different types of NPs functioning as subjects, and their distribution across clause types. We will compare the findings to Du Bois’s idea of preferred argument structure (Du Bois 1985, 1987a and 1987b).

Table 15, which is based on Tables 12–14, compares the different NP types (personal pronoun vs. other pronoun vs. full NP) and their distribution in the subject role across different clause types (transitive vs. intransitive vs. copular).

Table 15. Type of subject NP across clause type.

	St		Sc		Si	
	N	%	N	%	N	%
1st and 2nd pers. pron	103	52.3	19	9.6	34	21.4
3rd pers. pron, human ref	46	23.4	27	13.7	21	13.2
3rd pers. pron, non-human ref	12	6.1	119	60.4	54	34.0
full NP, human ref	32	16.2	12	6.1	12	7.5
full NP, non-human ref	4	2.0	20	10.2	38	23.9
Total	197	100	197	100	159	100

Table 15 shows the distribution of different types of subject NPs in transitive, intransitive and copular clauses.

Figure 1 presents the data from Table 15 as a graph. We can easily see that full NPs (the two rightmost columns) are a minority compared to pronominal subjects (the three leftmost columns). 1st and 2nd person pronouns function most often as transitive subjects, whereas 3rd person pronouns referring to non-human referents figure significantly in the intransitive subject and copular subject roles. There are not very many full NP subjects in transitive clauses, whereas in intransitive clauses, the percentage is considerably higher: as Table 15 shows, 18.3% of transitive subjects and 31.4% of intransitive subjects are full NPs.

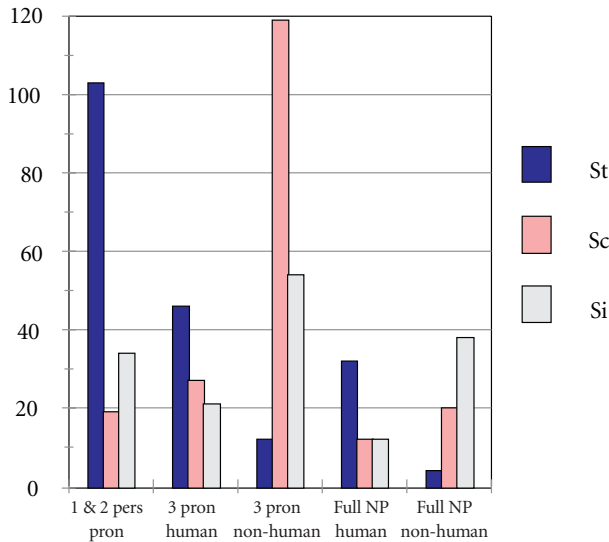


Figure 1. Type of subject and grammatical role.

The findings presented in Figure 1 support the claims of Du Bois (1985, 1987a) regarding preferred argument structure, especially the “Non-Lexical A Constraint” proposed by Du Bois, according to which there is an overwhelming tendency in spoken discourse for NPs in the transitive subject role (in Du Bois’s terminology, A) to be non-lexical. However, we can also see that pronominal subjects outnumber lexical NPs in intransitive and copular clauses as well (109 out of 159 intransitive subjects are pronominal (68.6%) and 165 out of 197 copular subjects are pronominal, 83.8%). In contrast, in the object role, full NPs outnumber pronominal objects (106 out of 237 objects are pronominal, i.e. 44.7%). Thus, the data do not show the kind of alignment of the intransitive subject with the object that Du Bois (1987a:822) found in his data, where almost half of both objects and intransitive subjects were lexical.

Table 15 and Figure 1 show that 1st and 2nd person pronouns figure significantly in the subject role: 1st and 2nd person pronouns make almost 40% of all subjects and the majority of transitive subjects. We can also see that the majority of non-human pronominal subjects (119/185; 64.3%) are subjects of copular clauses, whereas only a few function as transitive subjects (12/185; 6.3%). Looking at the same figures from the perspective of copular clauses, in 60% (119/197) of copular clauses the subject is a 3rd person pronoun with a

non-human referent. In intransitive clauses there is not such a clear preference, but even there, the largest number of subjects are 3rd person pronouns referring to non-humans (34%; 54/159).

Figure 1 is extremely interesting in light of our discussion of explicit and variable coding (Sections 2.3–2.6). Figure 1 shows that those NPs which exhibit the most explicit coding when functioning as subjects (i.e. personal pronouns) appear mostly in transitive clauses. In contrast, coding is less explicit in intransitive and copular clauses, since the majority of subjects in these clause types are either pronouns referring to non-humans or full NPs. Of course, the functional load on the codings to distinguish different functions is much greater in transitive clauses than in intransitives. Thus, we can see that personal pronouns are central in the subject role not only in terms of frequency but also, in terms of coding patterns: in personal pronouns, the subject role is coded most explicitly.

As we have seen, personal pronouns realize a straight-forward nominative-accusative coding pattern in terms of case marking, agreement, and word order. Regarding case marking, 1st and 2nd person subjects are marked with the nominative, and objects with the accusative or the partitive. In the agreement system, 1st and 2nd person subjects trigger explicit codings, and with regard to word order, personal pronouns exhibit a relatively fixed Subject–Verb ordering. Thus, personal pronouns exhibit clearly grammaticized codings of the subject role.

In contrast, the coding of full NPs in Finnish approaches a neutral coding system in some respects. Case marking of full NPs is not distinctive in the plural, nor is agreement, since oppositions are neutralized. In word order patterns, as I have shown, there is a cline from the relatively fixed Subject–Verb ordering of personal pronouns and 3rd person pronouns to the more flexible ordering principles of lexical NPs. Verb–Subject order is most frequent when the subject is a full NP referring to a non-human referent. Thus, the coding of full NPs does not clearly distinguish the subject as a distinct grammatical role, but instead, there is much overlap in the coding system. Nevertheless, full NPs can be recognized as subjects based on the model of pronominal subjects.

3.1.2 The definition of subject from a discourse perspective

In Finnish linguistics, the category of subject has traditionally been taken to include a fairly broad range of grammatical elements. Instead of a strict definition that would define a unified grammatical role, scholars have preferred to

include in the category of subject types of NPs which call for various specifications that are idiosyncratic to certain verbs and clause types. Many recent discussions acknowledge that there is a prototypical area in the category of subject that includes NPs which are in the nominative case and which trigger agreement in the verb, i.e. fall under the definition that I have given here for the subject, but in addition to these, there has been a desire to include marginal cases as well (see e.g. A. Hakulinen 1983, Vilkuna 1989, 1996). We will discuss one such case shortly (see below the section on E-NPs).

Interestingly enough, the marginal cases have received much more attention in the literature on the subject role in Finnish than the more central or prototypical subjects. This may be partly due to the fact that the literature has largely been based on example sentences taken from the linguistic literature rather than naturally occurring data. Even studies that have used naturally occurring data (such as Hakulinen and Karlsson 1979, Vilkuna 1989) are not data-driven, but rather, data are used to illustrate the theoretical claims that are being made.

In my work, I have tried to take the patternings found in the data as my starting point. I have looked at how the different coding strategies are put to use in the coding of grammatical relations. The subject role defined by these strategies takes the nominative case and triggers agreement. The data also show that the subject NPs thus defined almost always take a preverbal position. Furthermore, the data analysis highlights the prominent role of personal pronouns in the grammaticization of the subject role. In conclusion, with a narrower definition of “subject”, NPs that function as subjects represent a category that is more unified grammatically and also shows a distinct discourse profile, as I will show in the next section.

3.2 What do subjects do in discourse?

In this section, I discuss the use of subjects in discourse. I will start by discussing the relationship between the introduction of new referents and the subject role. I will then look at how referents are treated after they have been introduced, i.e. whether they tend to be tracked or dropped after the first mention. Then I will consider the subject role with respect to the semantic class of the referent. Finally, I will discuss the idea of subjects as “starting points” which open up a perspective from which the predication proceeds.

3.2.1 Information flow and the subject role

According to Chafe, intonation units typically contain only one piece of new information, and the new information is not likely to be carried by the subject constituent (Chafe 1979, 1994). Based on this principle, Chafe (1994) has proposed a “light subject constraint” according to which subjects tend to refer to referents that are either not new to the discourse or are new but of “trivial importance”. In this section, we will look at certain features of information flow in relation to the subject role and also, discuss the implications of these findings.

We have already seen in previous sections that pronominal NPs outnumber full NPs in the subject role; we would expect attenuated forms of reference such as pronouns to refer to referents that are not new. Table 16 illustrates the relationship between grammatical role and activation cost in the Finnish data.

Table 16. Activation cost and grammatical role.

	New		Total
	N	%	N
St	29	10.4	279
Sc	44	19.0	231
Si	44	20.6	214
O	113	38.0	297
Obl	186	40.8	456
E-NP	89	75.4	118
Free NP	35	34.3	102

Table 16 shows that subjects have low percentages of new mentions, whereas the percentages of new mentions in the object and oblique roles are considerably higher (around 40%). However, subjects of transitive, intransitive, and copular clauses do not behave alike with respect to activation cost, but instead, intransitive subjects and subjects of copular clauses are much more likely to carry new mentions than transitive subjects (around 20% of Si and Sc are new, but only 10% of St). Note also that the vast majority of E-NPs have referents that are new to the discourse (see next section for further discussion), whereas about one third of free NPs are used to bring new referents into the discourse.

Huumo (1994) reports on similar findings regarding the relationship between activation cost and the subject role in his study of Finnish literary fiction. In his data, 14.2% of the subjects were mentions of referents new to the discourse. In comparison, the percentage of new mentions in the subject role in my data (i.e. St, Si, and Sc) is 16.2%. However, the numbers are not fully comparable because of differences in the definition of subject: I have not included existential clauses among intransitives (for further discussion, see next section).

In his study of Sakapulteko spoken narratives, Du Bois (1985, 1987a, 1987b) found that there is an ergative patterning in his data in which the transitive subject role stands out as the locus for given information (only 3.4% of transitive subjects in the Sakapulteko data referred to new referents), whereas intransitive subjects align with objects in carrying more new information (32.8% of intransitive subjects and 23.7% of objects referred to new referents). To account for this patterning, Du Bois proposes the “Given A Constraint”. As we can see from Table 16, the Finnish data do conform to this constraint, but the data do not suggest any ergative patterning in managing new information; instead there is a continuum from transitive subjects, which contain the fewest new mentions, to E-NPs, with the most new mentions.

St > Sc > Si > Free NP > O > Obl >> E-NP

As can be seen in Table 16, there is a gap between the obliques and E-NPs in the percentage of new mentions, and the E-NP role stands out as the grammatical role which is most suitable for presenting new information. However, if we compare the absolute numbers, we can see that the oblique role carries the largest number of new mentions.

The continuum that can be seen in the proportion of new mentions in Table 16 offers one perspective on the management of new information in discourse. From another perspective we could relate introduction of new referents to how the referents are treated in subsequent discourse. As we saw in Table 10, there is an important difference between new mentions made in the core roles and new mentions in the oblique role: those referents that are introduced in the oblique role are hardly ever mentioned again in the discourse (18 referents of the total of 186 new mentions were mentioned again, i.e. 9.7%, see Table 10 above), whereas referents that are first mentioned with NPs in the core roles are much more likely to be tracked in the discourse (see above Section 2.4.3, and Durie 1994).

Thus, the core roles do important discourse work by keeping track of participants. Consider Table 17.

Table 17. Referent tracking and grammatical role.

	Tracking		Total
	N	%	N
St	250	89.6	279
Sc	205	88.7	231
Si	187	87.4	214
O	232	78.1	297
Obl	114	25.0	456

Table 17 shows that subjects most often have referents which are being tracked in the discourse. In the object role the percentage of tracking mentions is slightly lower. In the oblique role, the percentage is only around 25%. Thus, NPs in the subject role contain the most tracking mentions, and NPs in the object role somewhat fewer, whereas in the oblique role, only a minority of the NPs are tracking (see Thompson 1997 for similar findings for English conversation).

As was discussed in Section 2.6.2, transitive subjects typically refer to human referents. Table 18 illustrates the distribution of human referents across grammatical role.

Table 18. Distribution of human referents across grammatical role.

	Human		Total
	N	%	N
St	255	91.4	279
Sc	68	29.4	231
Si	98	42.4	214
O	19	6.4	297
Obl	91	20.0	456

From Table 18 we can see that the transitive subject role stands out as the locus for human referents, as over 90% of transitive subjects have human referents, whereas fewer than half of intransitive subjects do (about 42%). In copular clauses, fewer than 30% of the subjects are human. In the object role,

the referents are mainly non-human.¹ There is evidence from other languages that the low percentage of human referents might be characteristic of the object role even in more general terms (see Du Bois 1987a for Sakapulteko, Herring 1989 for Tamil, Ashby and Bentivoglio 1993 for French and Spanish, Nakayama and Ichihashi-Nakayama 1994 for Japanese, Kärkkäinen 1996 for English).

The distribution of human referents across grammatical role may not be so interesting per se, but it does give us some perspective on what NPs in the different grammatical roles can be used for in discourse. Transitive subjects often appear as agents that instigate a process or an event, and agents, of course, are typically human. In contrast, copular clauses are most often being used to make assessments about the referent of the subject NP or to identify it. It is interesting to find then that, at least in the Finnish data, in copular clauses assessments are most often made about referents that are not human. Tainio (1993:83) offers supporting findings, also based on Finnish conversational data, noting that the subject-NPs in copular clauses most often refer to 2nd order entities, i.e. to events, processes, and states-of-affairs (see Lyons 1977:443).

In conclusion, we have seen that the information flow features discussed in this section characterize the subject role in interesting ways. First, new referents are rarely introduced in the subject role. Instead, there are other grammatical roles that are much better suited for this purpose, most notably the E-NP (for more discussion, see below Section 3.3). We also noted that there is a difference between new mentions made in the core roles and new mentions in the oblique role: those referents that are introduced in the oblique role are hardly ever mentioned again in subsequent discourse, whereas referents that are first mentioned in the core roles are much more likely to be mentioned again, i.e. to be tracked in the discourse. Finally, we looked at the distribution of human referents across grammatical roles, and we saw that the subject role is much more likely to contain mentions of human referents than the object role. In sum, the subject function is not so much used for bringing new referents into the discourse but instead, it is well suited for keeping track of referents. Among the core roles, transitive subjects are mainly used for tracking human referents, while intransitive subjects accommodate other kinds of referents also. Thus,

1. The percentage of human referents in the oblique role may seem surprisingly high. As was discussed in Section 2.4.5 on the adessive, adessive NPs in possessive constructions are in many respects different from other oblique, *inter alia*, in preferring human referents. If we take out adessive NPs from other obliques in Table 17, the percentage of human referents in the oblique role drops to 10%.

there are several information flow features that characterize NPs in the subject role and distinguish them from NPs in other grammatical roles.

3.2.2 Split patterning and the grammaticization of the subject role

As was discussed in Section 2.2, it is common across languages that coding systems are not unified, but instead, exhibit split patterns. Splits are especially common in ergative languages. Silverstein (1976, 1981) has proposed a hierarchy of “inherent lexical content” to account for split patterns in ergative languages. We will now discuss this hierarchy and its relation to the coding of the subject role in Finnish.

In Silverstein’s hierarchy, personal pronouns which refer to the speech act participants rank highest. They are followed by third person pronouns, proper nouns and kin terms, nouns that refer to humans, animates, and inanimates (Silverstein 1976: 122). According to Silverstein, if there is a split in an ergative coding system, the hierarchy of “inherent lexical context” predicts where the split should occur: those referents that are ranked highest in the hierarchy are more likely to receive nominative-accusative coding, whereas referents at the lower end of the hierarchy are coded according to an ergative pattern. For example, if in a given language, third person pronouns receive nominative-accusative coding, then items that are higher on the hierarchy, i.e. 1st and 2nd person pronouns, are marked according to this pattern as well.

Although Finnish is not an ergative language, Silverstein’s hierarchy provides valuable insights regarding the coding of grammatical relations in Finnish. As we have seen in the previous sections, personal pronouns in Finnish exhibit a clear nominative-accusative system in terms of case marking, agreement, and word order. In contrast, the coding of full NPs in Finnish is close to a neutral coding system. In word order patterns, personal pronouns show a relatively fixed Subject–Verb ordering, whereas lexical NPs show more flexible ordering principles. As we saw in Section 2.6.2, Verb–Subject order is most frequent when the subject is a full NP with a non-human referent.

This split between explicit and variable coding is in accordance with the Silversteinian hierarchy, since the nominative-accusative coding system is realized most explicitly in the coding of referents that are highest in the hierarchy, namely, 1st and 2nd person. Since the nominative-accusative coding system is designed to code the subject role, we can say that the coding of 1st and 2nd person pronouns shows a highly grammaticized subject role, whereas the coding of full NPs does not provide such strong evidence for a grammaticized subject role.

According to Silverstein (1976: 113), the speech act participants (expressed by 1st and 2nd person pronouns) are the most natural agents, grammatically expressed by transitive subjects, and the least natural patients, i.e. grammatical objects. He proposes that there is a “metapragmatically based understanding of predication-perspective” according to which “language structure is grammaticalising (with characteristic asymmetry) the perspective from which a state of affairs is predicated of referents, the most ‘natural’ being that which grows out of the configuration of the ongoing speech event, the informational givens of which, the maximally presupposable entities, are of course of the same ordering as above [i.e. in the hierarchy].” (Silverstein 1981: 243).

Applying Silverstein’s ideas to our framework, we could say that grammatical coding offers a grammaticized expression to the perspective from which speakers make predications about referents. According to Silverstein, the most “natural” perspective is one which grows out of the current discourse context (Silverstein’s “ongoing speech event”), especially the point of view of the speech act participants. However, it is not entirely clear what Silverstein means with “natural” here, whether it means natural from a semantic point of view or from the point of view of language use.

In some sense, Silverstein’s “predication-perspective” comes very close Chafe’s idea that the subject role gives a grammaticized expression to the “starting point” function. Chafe (1994: 83) has proposed that “each [clause] has a point of departure, a referent from which it moves to provide its own new contribution. It is this starting point referent that appears grammatically as the clause’s subject.” (See also Chafe 1976.) This is extremely interesting in view of the fact that so many subjects are 1st or 2nd person pronouns: the speech act participants are quite natural starting points in interaction. In a very concrete way, we “move” from the speech act participants “to provide new contributions”. If we think about the subject as grammaticizing the starting point function, it is not surprising to find that 1st and 2nd person pronouns exhibit a straight-forward nominative-accusative coding system, and accordingly, a highly grammaticized subject role.

When discussing the grammaticization of the subject role in Finnish it is important to bear in mind that in Finnish, subjectless clauses are perfectly acceptable. In subjectless clauses, there is usually some constituent — most often a locative or temporal element — before the verb, i.e. in the normal position for subjects (see Vilkkuna 1989: 40–50). Consider the following example, which comes from a conversation where the participants are talking about Linnea’s summer cottage.

(34) Women

Linnea: *siit mei-l on nyt uus uima-paikka*
 it+ELA 1PL-ADE is now new swimming-place
teh-ty sinne.
 make-PSS+PCP there
 There we have now made a new swimming place there.

Example (34) contains a blend of two constructions: it starts out as a possessive construction (see previous chapter on the adessive) *meil on nyt uus uimapaikka* ‘we have a new swimming place’, but then it is reanalyzed as a passive *uus uimapaikka (on) tehty* ‘a new swimming place has been made’. In either construction, there is no subject (for a discussion of the object in passive clauses, see Chapter 1). Instead, there is a locative element before the verb, or actually two, the pure locative *siit* ‘from there’ and the possessive *meil* ‘on us/to us’. This choice of grammatical constructions is interesting: as we saw in Section 2.5 on agreement, the passive in Finnish is a personal form, but if there is no 1st person plural pronoun with the passive, the reference is not specified but has to be inferred from the context. In example (34), the reference of the passive person marking can be interpreted at least in two ways: it is either co-referential with the referent of *meil*, the locative constituent in the possessive construction, or the reference is open, with the interpretation of the whole construction as meaning that the referents of the ‘we’ (*meil*) had the swimming spot made for them. In some construction types, inter alia in existential and possessive constructions, the locative constituent has been grammaticized to the preverbal position (see Huumo 1995). Thus, the starting point function does not seem to characterize the subject role distinctively, but could rather be understood as a more general organizing principle of clause structure.

In the previous sections I have shown that there is a split between explicit and variable coding in the expression of grammatical relations in Finnish. In this section, I showed that the NPs that exhibit most explicit coding patterns, namely 1st and 2nd person pronouns, appear most often as transitive subjects. But I also showed that they make up to 40% of all subjects. We also discussed the idea proposed by Chafe (1994) that subjects serve as grammaticized expressions of the starting point of the clause that opens up a perspective from which the predication proceeds (cf. Silverstein 1976 on “predication-perspective”).

3.3 Subject and non-subject

In our discussion of the coding of the subject, I have argued that only those NPs which are in the nominative and which trigger agreement should be considered subjects in Finnish. In this section, I would like to search for further support for this definition by contrasting it with two types of NPs which have traditionally been considered to be subjects in Finnish linguistics, namely the so-called existential “subjects” and “subjects” with open quantification.

Our definition of the subject role has been based on morphosyntactic codings: it makes reference to case (nominative) and agreement patterns (an NP functions as a subject if it triggers agreement in the verb). I have shown that the subject thus defined shows a distinctive discourse profile (subject NPs are mainly used to track referents, but rarely used to bring new referents). The definition is fairly abstract and contains no information that would be specific to certain sets of verbs or clause types. As I will show in the next subsections, “subjects” of existential clauses and “subjects” that are partially quantified fall outside the category of subject because they only partially satisfy the criteria for subjects and, moreover, they are characterized by several features that are idiosyncratic to a certain set of verbs or to a clause type.

3.3.1 Existential constructions

As was discussed in Section 2.4.5 on the adessive, Finnish has a special existential construction that was illustrated with the following schema (here repeated as 35):

(35) Existential clause schema

$\text{NP}_{\text{LOC/POSS}} + \textit{olla}$ ‘be’ + $\text{NP}_{\text{NOMPTV}}$

The same schema is used in possessive constructions. In the possessive construction, the locative NP is usually in the adessive, but in the other existentials it may be in some other locative case. It is quite common across languages for spatial constructions to be extended to the expression of possessive relationships as well (see e.g. Lyons 1977, Heine et al. 1991, Heine 1997).

In the existential construction, it is possible to use some other verbs with an existential meaning in addition to the copular verb *olla* ‘to be’. However, the set of possible verbs is fairly small (Wähämäki 1975). The postverbal NP may be either in the nominative or in the partitive case. I have referred to this NP as the E-NP (the so-called “existential subject” in traditional Finnish linguistics; for a more detailed discussion, see Helasvuo 1996). Existential constructions exhibit

a relatively fixed word order with the locative NP preceding and the E-NP following the predicate.

Example (36) illustrates how existential constructions are used in discourse. The excerpt comes from the conversation between four men who are at this point discussing their kids and how they wear out their clothes so quickly.

(36) Alko

- 1 Aarre: *kyllä meillä vesseli kuluttaa et,*
Sure in our family the little rascal wears out (his clothes) so that,
- 2 *ei likka kuluta kyllä niin kauheesti.*
the girl does not wear out (her clothes) so quickly.
- 3 *poika kyl [1lä helkkari1] se ku,*
The boy oh damn it, he...
- 4 Pekka: *[1joo joo.1]*
Sure, sure.
- 5 *[2ne on nii.2]*
They are so,
- 6 Aarre: *[2ei si-l täy2] tyy ol-la semmose-t,* ←
NEG+3SG 3SG-ADE must be-INF a.kind.of-PL
No, he has to have these,
- 7 *oikeen si-l on,* ←
really 3SG-ADE be+3SG
he even has
- 8 *...(0.3) nyt semmose-t nah [3ka-housut jala-ssa ja.3]* ←
now a.kind.of-PL leather-pants leg-INE and
now these leather pants on and.
- 9 Pekka: *[3nii-t on erilais-i-a3]*
3PL-PTV is different-PL-PTV
They have different kinds of
- 10 *taipumuks-i-a nii-llä [4ja.4]*
inclination-PL-PTV 3PL-ADE and
inclinations they do and...
- 11 Aarre: *[4ja sit4] si-l on* ←
and then 3SG-ADE be+3G
and then he has

- 12 *viel kura-housut,*
also mud-pants
rubber pants also,
- 13 @@*että semmose-t et* $\langle X_{[5}X \rangle_{\cdot 5}]$
that a.kind.of-PL that
so that (he has) these — .
- 14 Tor: $[_5 \textit{jaa ne on}_5]$ *sellase-t*
oh 3PL be+3SG a.kind.of-PL
Oh they are kind of
- 15 *niinku* $[_6 \textit{muovise-t}_6]$
like plastic-PL
like plastic.
- 16 Aarre: $[_6 \textit{ei si-llä}_6]$ *muu-t kest-ä.*
NEG+3SG 3SG-ADE other-PL last-INF
other (kinds of clothes) wouldn't last for him.

In example (36), Aarre starts out by comparing his son and daughter (lines 1–3). Pekka responds to this by starting a predicate nominal construction which is not completed (line 5). Overlapping with Pekka, Aarre continues about his son (lines 6–8). He characterizes him as a little rascal by describing his clothing (lines 6–8, 11–12, 16). These characterizations are made using existential constructions *sil on semmoset nahkahousut* ‘he has these leather pants (lit. on him are these leather pants)’ (lines 7–8) and *sil on kurahousut* ‘he has rubber pants’ (lines 11–12).

On the basis of constructed and anecdotal evidence, Wähämäki (1975) has suggested that the discourse function of existential constructions is to introduce new referents. As we saw in Table 16, this is true of conversational language also: E-NPs are most often used to bring new referents into discourse (75% of E-NPs refer to referents that are new to the discourse), as exemplified in example (36): E-NPs introduce new referents into the discourse (*nahkahousut* ‘leather pants’, line 8, and *kurahousut* ‘rubber pants’, line 12). After the first mention, the referent of *nahkahousut* is never mentioned again and the referent of *kurahousut* is mentioned once (*ne* ‘they’, line 14) and then dropped. This example illustrates a more general pattern: referents that are first introduced in the E-NP role tend not to be tracked, but instead, they are dropped after the first mention. Specifically, only 26 of the 89 referents that are introduced as E-NPs are ever mentioned again (i.e. 29%). As was noted above, low frequency of further mentions is characteristic of oblique arguments. Thus, although

E-NPs are used to introduce new referents, these referents are mostly not kept track of, but instead they are dropped.

In an existential clause, the new NP is presented in relation to a spatial or possessive frame that is expressed by the preverbal locative NP. From a discourse point of view, it seems that it is the frame that is most relevant to the discourse, not the presented referent. In example (36), the frame is the possessor, expressed by the 3rd person pronoun *se* (in the adessive case: *sil(lä)*). In this sequence, it is the referent of the locative NP, namely Aarre's son, that is being tracked. The referent is being characterized as a little rascal by describing his clothing.

In the existential construction, the frame serves as the starting point of the clause (cf. the previous section). Huumo (1996b: 297) proposes that existential clauses function to state “a predication about the locative space by reporting its content”. I would like to argue that existential clauses function to characterize the locative space by establishing a relationship between it and a new referent. The locative space may be either concrete or abstract.

In Finnish linguistics, the E-NP has traditionally been analyzed as a subject, although it has been admitted that it is atypical in many ways (see e.g. Hakulinen and Karlsson 1979, Vilkuna 1989 and 1996). It does have some of the features typically associated with subjects. Namely, it is the first argument of the verb, given that it is the only NP that is in a non-locative case. However, the E-NP lacks most of the subject coding features discussed above. It occurs usually after the verb and does not trigger agreement. It may be in the nominative, but also in the partitive. Further, E-NPs tend to be full NPs ((97/118; 82.2%), while most subjects are pronominal. The referents of E-NPs are mostly new to the discourse (89/118; 75.4%) and moreover, they are usually non-human (108/118; 91.5%). Comparing this to Tables 17 and 18, we can see a striking contrast with the profile that NPs in the subject role show, as subject NPs rarely bring new referents into discourse. Further, over 90% of the referents of transitive subjects are human.

As was noted in the previous section, however, subjects of copular clauses tend to have non-human referents, as only about 30% of copular subjects are human (Table 17). Both existential and predicate nominal clauses serve to make characterizations of referents (see example 36, where existential constructions alternate with copular clauses in making characterizations of referents, lines 5, 9–10, 14–15). However, in predicate nominal clauses, the characterizing NP is often non-referential, whereas in existential constructions the E-NP is referential. Thus, in an existential construction, the characterization is made in a different

way, namely, by establishing a relationship between two referents. (On the similarities between existential clauses and predicate nominal clauses, see Toivainen 1986, and Helasvuoto 1996a.)

Looking at the existential constructions from a grammaticization perspective, we can see that the grammaticization of the E-NP is totally different from that of the subject role. As I have shown in the previous sections, it is possible to describe the grammaticization of the subject in Finnish in fairly abstract terms without making reference to specific constructions or lexical restrictions. In contrast, the grammaticization of the E-NP involves the grammaticization of the existential construction as a syntactic schema. In an existential construction, the case marking and position of the E-NP have been crystallized as well as the choice of the verb (*olla* 'to be' and certain other verbs) and the position and case marking of the locative NP.

In conclusion, E-NPs do not have the coding features that characterize subjects in Finnish. They do illustrate, however, how the function of subject is built up from several features that together work to make subjects useful in grammar and in discourse for certain purposes, but perhaps not for others. Inter alia, the subject function is not well suited for introducing new referents into discourse. This is reflected in the word order patterns: the postverbal position is the preferred locus for new information in Finnish, and since subjects rarely refer to referents that are new to the discourse, they are predominantly preverbal. Furthermore, subject NPs are mainly pronominal. In contrast, the E-NPs offer a grammatical slot for introducing new non-human referents. Consequently, they are almost always full NPs and take a postverbal position.

3.3.2 Subjects and open quantification

As was discussed in Section 2.4.3, the partitive is mainly used to mark objects. In the marking of the object, the partitive can be used to indicate open quantification (Heinämäki 1983, Larjavaara 1988).² The partitive can also be used to mark open quantification in E-NPs, and it is sometimes also used in connection with verbs that do not merely indicate existence or coming into existence, but something more, such as motion. In this section, we will discuss the use of the partitive in open quantification of NPs other than objects.

2. Heinämäki (1983, 1994) speaks about boundedness (see also Leino 1991): the choice of the accusative as an object marker indicates boundedness, whereas the partitive indicates unboundedness.

Consider example (37), which comes from a conversation between two brothers, Pekka and Simo. Pekka has been in Northern Finland fishing and skiing and is telling his brother about it:

(37) Brothers

- | | | |
|---|--------|---|
| 1 | Pekka: | <i>no sit alko tule-maa iso-i maki-i</i>
oh then start+PSS+3SG come-INF big-PL+PTV hill-PL+PTV |
| 2 | | <i>eteen ja,</i>
in.front and
Well then there appeared (lit. started to come) big hills
ahead of us and... |

The construction that is used in example (37) is based on an anthropocentric metaphor which conceptualizes the person as a static center towards which the surroundings move. In the example, the big hills started to appear in front of the deictic center, i.e. the speaker. The deictic center is not expressed at all grammatically, but rather, the postposition *eteen* ‘in front (of)’ is used without an NP expressing whose front is being talked about; rather, it has to be inferred. The big hills are visualized as starting to appear or approach the inferred center, but the use of the partitive *iso-i maki-i* ‘big-PTV hills-PTV’ leaves the quantity of the hills open.

In terms of discourse frequency, partitive NPs that do not function as objects are peripheral: in my data, there are 126 partitive NPs that function as objects, whereas there are only 49 non-object partitives, of which 45 are in existential or possessive constructions (E-NPs). This leaves only 4 partitive NPs that are neither E-NPs nor objects, but arguments of intransitive verbs.

The question is, then, whether the partitive NP (such as *iso-i maki-i* in example 37) should be analyzed as a subject or not. It bears some resemblance to the subject: *inter alia*, it is the first argument of a verb that usually takes a subject. However, the set of verbs with which it is possible to use the partitive is very limited (the copula or certain intransitive verbs; for more discussion, see Wähämäki 1975). Furthermore, the verb does not show agreement with the partitive NP: in example (37), the partitive NP is in the plural but the verb is in the singular. As we have seen, the number opposition has been neutralized to a great extent in the spoken language. Here, however, number agreement would not be possible even in the standard written language. Thus, agreement does not code partitive NPs as subjects. The information flow features that characterize partitive NPs are atypical of subjects. Namely, partitive NPs most often refer to non-human referents that are new to the discourse. As we have seen in the

previous sections, these features correlate with the absence of coding. Thus, there is very little support to an analysis of partitive NPs as subjects.

3.4 Summary on the grammaticization of the subject

We have discussed the grammaticization of the subject role both in terms of grammatical coding patterns and in terms of information flow features. I have shown that by defining the subject fairly narrowly, the resulting grammatical category is characterized by several grammatical and discourse features that are distinctive for that category. In contrast, the NP types that we have discussed in the last two subsections, namely “subjects” of existential clauses and “subjects” that are partially quantified, fall outside the category of subject because they only partially satisfy the criteria for subjects and, moreover, show a discourse profile that is very different from that of the subject. In their grammaticization pathways, these constructions are characterized by several features that are idiosyncratic to a certain set of verbs or to a specific construction type.

I have shown that E-NPs and partitive “subjects” lack most of the coding features that characterize subjects: they do not trigger agreement, and E-NPs may appear in the partitive case as well as in the nominative. Regarding word order, E-NPs are always in postverbal position. We have seen that subjects are most often pronominal, whereas the majority of E-NPs and “partitive subjects” are full NPs. In terms of the discourse work NPs do, E-NPs are used for introducing new non-human referents. In subsequent discourse, however, these referents are not usually tracked, but are instead dropped. In contrast, subjects serve to track participants that have already been introduced to the discourse. The referents of subject NPs are usually human. Thus, their respective discourse profiles are completely different, and so are the coding features. We can conclude that they do not represent the same syntactic function, i.e. the subject.

We have discussed the subject role from a grammaticization viewpoint. I have shown that personal pronouns exhibit highly grammaticized nominative-accusative codings, and accordingly, a clearly defined subject role. In contrast, the marking of full NPs is closer to a neutral alignment system. Accordingly, the coding does not clearly distinguish the subject as a distinct grammatical role, but instead, there is much overlap in the coding system. We have also seen that in terms of actual frequencies, personal pronouns figure significantly in the subject role in spoken discourse, whereas full NP subjects are a small minority.

CHAPTER 4

Free NPs

In the section on agreement (§2.5), it was noted that in Finnish there is agreement not only between the subject and the predicate, but also, within the noun phrase between the head and its modifiers. This is what Lehmann (1988) calls “internal agreement”. In Finnish noun phrases, nearly all determiners and modifiers precede their heads and agree with them in case and number. This is one of the defining features of noun phrase formation in Finnish. Since the determiners and modifiers already show the case and number of the whole construction, they also indicate the syntactic role the construction serves in the clause. We could say then that noun phrases in Finnish show a tight internal syntactic bond. In the preceding chapters, we looked at how noun phrases are used as clausal constituents, and in this chapter, we will look at the use of noun phrases as free NPs, i.e. NPs which are not part of any clause but instead form free constructions. I will show that there are several types of free NPs and that each type serves specific functions in discourse.

Hopper and Thompson (1984) discuss the distinction between nouns and verbs in terms of their prototypical functions in discourse. According to them (1984:708), the most prototypical function of nouns is to introduce a participant into the discourse, whereas prototypical verbs serve to assert the occurrence of an event. As we have seen in Chapter 2, nouns not only serve to introduce participants, but also, to track them in discourse. I will show below that free NPs can be used in these functions, i.e. in manipulating referents, but I will also demonstrate that they can be used to make predications.

I will first try to explicate what I mean by free NPs. Then I will discuss free NPs as a syntactic type in the light of the existing literature. We will then look at case marking of free NPs in Finnish, and finally, there will be a discussion of the types of free NPs with some examples.

4.1 What is a free NP?

Free NPs are noun phrases that are not part of any clausal construction, but instead, form free constructions. In this section, I will discuss the analysis in the light of examples from the data.

Those NPs that were coded as “free” in the data were syntactically clearly free and could not be analyzed as part of any clause in the context. However, this is not always an easy decision to make. Because the status of free NPs is not widely recognized, I have analyzed NPs as clausal constituents, whenever possible. If this has not been possible, I have analyzed the NPs as free. In the data, there were 102 free NPs thus defined. However, the coding imposes a binary distinction (clausal constituent vs. free NP) where there is no sharp line, but rather, a continuum of integration (see also discussion in Section 2.1.2 above).

The following example illustrates this. It comes from the Birthday party conversation. Raita, the primary speaker, is in the middle of a story of a cat who has committed suicide. Leena is acting as a co-teller. (For a more detailed discussion of the story, see Section 5.2)

(38) Birthday

- 1 Raita: [1 *se oli* 1]
It was —
- 2 Leena: [1 *mutta=* 1] *se oli*
but 3SG was
but it was
- 3 *eik se ollu Onni ja Pirkko ajanu se-n s-kissa-n*
NEG-3SG+Q it been Onni and Pirkko driven 3SG-ACC cat-ACC
wasn't it Onni and Pirkko who had driven the th- cat
- 4 *sii-hen tilantee* [2-*seen*.2]
3SG-ILL situation-ILL
into that situation.
- 5 Raita: [2 *niin*.2]
Right.
- 6 Raita: *neljäs tai viides kerros*,
fourth or fifth floor
fourth or fifth floor,
- 7 *kumpi se oli*.
which.one 3SG was
which one was it.
- 8 Leena: (TSK) ^ *Viron-kadu*-[3 *lla*.3]
Viro-street-ADE ←
At Viro street.

- 9 (X): [₃ ^*nii*.₃]
Right.

- 10 Raita: *joo Viron-kadu-lle sitte niinku,* ←
 yeah Viro-street-ALL then like
 Yeah. (On)to the Viro street then like...

In the example there is an oblique NP *siihen tilanteeseen* ‘into that situation’ in line 4 which is clearly integrated into the predication in line 3: the predication would not make sense without the oblique NP. In lines 6 and 7, Raita uses a free NP to make a predication about the apartment where the cat lived: it was the 4th or 5th floor. Prior to this, the apartment has not been referred to. In line 8, Leena adds more specific information with an adessive NP *Vironkadulla* ‘at Viro street’. The adessive NP can be analyzed as locating the apartment, or alternatively, giving the location of the whole event being described. Raita acknowledges Leena’s contribution with the particle *joo* ‘yeah’ and repeats the same NP but in a different case form, the allative *Vironkadulle* ‘onto Viro street’ (line 10). The allative NP is quite different from the preceding one in that instead of giving information about the surroundings, it serves to make a predication of its own about the cat. So far, it has only been mentioned that the cat committed suicide, and that the apartment was on the 4th or 5th floor. The allative NP in line 10 makes a predication about a movement onto something, and the assumed theme of the movement is the cat (see Section 4.4.2). In the context, there is no verbal predicate that would describe the movement, but from several pieces of information given with the case-marked NPs we can infer that the cat jumped to the street rather than, say, used the stairs to run down.

In example (38), the oblique NPs in lines 8 and 10 both express locations, and neither of them occurs in the same intonation unit with a predicate verb. The adessive NP in line 8 serves to add more specific information about the location of the incident being described and as such, it is part of the conversational activity that has been going on for some time at this point, namely that of giving more background information of the circumstances leading to the suicide. In contrast, the allative NP in line 10 adds a contribution of its own in making a predication about the cat and how it died. Thus, the three oblique NPs in the example vary in their degree of integration: The first one *siihen tilanteeseen* ‘into that situation’ is integrated into the clausal structure preceding it in the same intonation unit, whereas the second one, the adessive NP *Vironkadulla* ‘at Viro street’, is expanding on the on-going activity of the preceding talk but it is not syntactically integrated to any particular predication in the

context, and thus, is a free NP. The third one, the allative NP *Vironkadulle* ‘onto Viro street’ functions to make a contribution of its own, which is syntactically unattached.

In sum, the differences between NPs that are part of a clausal structure on the one hand and free NPs on the other may be quite subtle. In order to arrive at an understanding how the discourse is being structured syntactically, we need a close microanalysis on how the interaction proceeds.

4.2 Free NPs as syntactic units: A cross-linguistic perspective

Free NPs have not been discussed in the literature as a syntactic unit type until quite recently (see e.g. Helasvuo 1988, 1991 on Finnish; Tao 1992 and Ono & Thompson 1994 on English; Durie 1994 on Acehnese; Tao 1996 on Chinese; cf. also Bentivoglio 1992 on Spanish, who names free NPs as syntactic constructions but does not discuss them in any depth). There is a richer literature on one subtype of free NPs, the so-called dislocations or topic- and antitopic constructions (see e.g. Lambrecht 1981, 1988; Ashby 1988, Geluykens 1987, 1988, 1992). In some studies free NPs are discussed as a special subtype of the clause that is defective in that it does not have a verb. I will first discuss two such studies in order to show that such approaches do not do justice to the nature of the phenomenon. Then I will turn to the functions of free NPs from a cross-linguistic perspective.

Halliday (1985) discusses phenomena related to free NPs as “minor clauses”, clauses which do not have a “mood structure”, i.e. they cannot be divided into subject and “finite” (close to our term predicate). According to Halliday (1985:63), minor clauses typically function as calls, greetings and exclamations, such as *Mary! Good night! Well done!* As we can see, the concept of minor clauses includes some constructions that we would call free NPs (the first two examples; the third one is not a NP, and hence, falls outside the scope of “free NPs”). Not included are titles and labels, although Halliday admits that they have some features in common with minor clauses (e.g., they lack thematic and mood structure). He notes (1985:63) that titles and labels should not be regarded as clauses “because they have no independent speech function”. Unfortunately he does not specify what he means by this. What Halliday seems to be proposing then is that a vocative, such as *Mary!*, is a clause (a minor clause), because it has an independent speech function, but *Nixon off the record*

(Newsweek 11/3/97) is not a clause, because it does not, according to Halliday, have an independent speech function.

In my view, Halliday is exactly right in saying that titles are not clauses — not for the reason he gives (that they do not have an independent speech function), but rather, because they do not exhibit the kind of grammatical structuring that is characteristic of clauses. Likewise I think that *Mary!*, used as a vocative, should not be regarded as a clause because it lacks clausal grammatical structure.

In a study on information flow in Tamil, Herring (1989) discusses free NPs that serve to introduce a new referent into discourse.

(39) (Taken from Herring 1989: 123)

Aṅkiṭṭē iruntu, oru kuṇa keḷavi.
 there from one hunchbacked old.lady
Anta keḷavi taṇṇi eṭuttukiṭṭe varṛā.
 that old.lady water carrying come-PR-3FEM.SG
 From the other direction, a hunchbacked old lady.
 That old lady comes carrying water.

In example (39), there is an NP *oru kuṇa keḷavi* ‘a hunchbacked old lady’, together with a locative phrase *aṅkiṭṭē iruntu* ‘there from’. The locative + free NP construction is used to introduce a new referent. In the following clause, a predication is made concerning this referent. Herring (1989: 137) calls this construction type “verbless presentation”, and mentions that there are several other construction types in Tamil which likewise “lack any surface predication whatsoever” (Herring 1989: 125). She calls these constructions “verbless sentences” (1989: 137).

According to Herring, the main function of verbless presentations is to introduce new referents, since almost 90% of verbless presentations bring a new referent into discourse (1989: 128). She takes this finding to cause serious problems for the idea of Preferred Argument Structure as proposed by Du Bois (1985, 1987a), especially for Du Bois’ proposal that the object and the intransitive subject role provide loci for new information (see Chapter 3 for discussion). Herring shows that verbless presentations are even more specialized for presenting new information than are objects and intransitive subjects. What Herring fails to see is that verbless presentations do not have anything to do with Preferred Argument Structure because they are not clauses and thus, do not have argument structure. Instead, verbless presentations are free NPs that serve to present topics (see Section 4.4.3 for discussion). As such, they do play

a role in how new information is managed in discourse. Du Bois proposed that underlying “Preferred Argument Structure”, there is a “Preferred Information Flow”, i.e. there exist “well defined patterns for the introduction of new information and the management of old information throughout a discourse” (Du Bois 1985:349). As I see it, Herring’s findings are not relevant to the study of Preferred Argument Structure, but rather, to the study of preferred patterns for managing information in discourse. Thus, Herring fails to see the full implications of her study because she seems to assume that the clause (or the sentence, as she calls it) is the only kind of syntactic unit that is relevant.

Herring (1989) is not the only one to have noted that free NPs can be used to introduce new referents into discourse. In a study of free NPs in English, Tao (1992) points out that they can function to introduce new referents. He also shows, however, that free NPs can be used by the co-participants (“respondents” in Tao’s terminology) in negotiating the identification of referents, such as acknowledging referent identification to the speaker who has introduced a referent, or requesting verification or confirmation of a referent introduced by the speaker. Consider example (40).

(40) (Taken from Tao 1992)

R: *I was talking to ... a gal at work,*

... ‘Lisa ... ^Green? ←

...(0.9) ‘Lisa ^Smith ‘Green?

...(1.3) *who’s going to have a ^baby-*

...(1.3)[Sh-] — -

L: [Does she have] ‘many ^already?

R: ... ^M- m.

In example (40), both arrow-marked lines contain a noun phrase, produced with a yes-no question intonation contour, which is characteristic of “try-markers” (Sacks and Schegloff 1979). Speaker R is checking whether the co-participant L can figure out who of the people that he knows is being referred to. R does not receive an immediate response, but only after a characterizing relative clause. L implies she knows the person with her question about the referent. Thus, Tao (1992) places introduction of new referents in the broader function of negotiating reference (see also Geluykens 1988, 1992 and Section 4.4 below).

Durie (forthcoming: Figure 3) shows that in Acehnese over half of free NPs are used to introduce new referents. In the Acehnese data, the new mentions that were made using free NPs resembled new mentions made in the core roles

in that they were “persistent”, i.e. the referents received further mentions after their introduction (Durie forthcoming: Table 9). This was true of my Finnish data also: in the data, there were 35 free NPs that contained mentions of new referents (out of a total of 102 free NPs), and 15 of these new referents were mentioned again (42.9%; compare Table 16 which looks at the distribution of new mentions across grammatical role and Table 10 which looks at case marking and further mentions of new referents).

Thus, in the Finnish data those new referents that are introduced with free NPs are quite often mentioned again, i.e. they are relatively “persistent”, to use Durie’s terminology. However, free NPs in Finnish cannot be considered to be grammatical constructions that would be specialized in introducing new referents, as only 34% of free NPs have referents new to the discourse. Furthermore, referential NPs are only one subtype of free NPs: in addition to their referential function, free NPs can be used to make predications (see Helasvuo 1988, 1991; Ono & Thompson 1994).

Ono & Thompson (1994) divide the functions of free NPs into two main groups, referential and predicating. They note that the predicating function has not received much attention in the literature (see, however, Helasvuo 1988, 1991), whereas free NPs serving referential functions have been discussed in several studies (most notably, in the literature on so-called dislocations; see Section 4.4.2 for references). This may be a matter of viewpoint also: some researchers have preferred to analyze predicating free NPs as clauses (see e.g. Longacre 1987). At least in the present framework of analysis, this approach to predicating free NPs seems quite far-fetched. The main reason why I have chosen to analyze predicating free NPs as free constructions and not as clauses is that they lack clausal argument structure. Although there is a predicate, namely the NP, it does not build an argument structure around it. I will come back to this point in Section 4.4.1.

Free NPs can also be used as vocatives. In terms of syntax, vocatives behave like other free NPs in that they are not integrated to clausal syntax. However, they may be different from other referential free NPs in terms of information flow features. For example, Herring (1989: 133) points out that in her Tamil data, vocatives never constituted a first mention, and thus, were always given, whereas other free NPs in her data (the so-called “verbless presentations”) provided a major way of presenting new information. However, in my data base there was only one vocative expression. My impression is that vocatives are not used very much in Finnish conversational discourse in general, and I have not included them in my discussion of the functions of free NPs (Section 4.4).

Given the fact that Finnish has rich nominal morphology both in terms of derivation and inflection, one might expect extensive use of NPs as free constructions. Kangasmaa-Minn (1984) discusses the division of labor between nouns and verbs in the Finno-Ugric languages and notes that they all exhibit some nominal case marking, and some, including Finnish, are characterized by a rich case marking system. In contrast to nouns, verbal markings are far less elaborate. Kangasmaa-Minn goes as far as to propose that the Finno-Ugric languages are *noun languages* as opposed to languages in which the verb is more important. According to her (1984:77), “the verb merely confirms what the nouns express.” It is not quite clear, however, what exactly Kangasmaa-Minn is comparing Finnish with. At least in the use of free NPs Finnish does not seem very different from other languages in which free NPs have been studied (Mandarin, Acehnese, English, Spanish).

There is cross-linguistic evidence for free NPs from languages which are typologically very different from Finnish, such as Mandarin (Tao 1996), Acehnese (Durie 1994), and English (Tao 1992, Ono & Thompson 1994). Inter alia, in the conversational English data of Helasvuo (1997a), 5.6% of intonation units contained free NPs (57 free NPs in 1008 intonation units), whereas in the present data, the percentage is 5.1 (102 free NPs in 1981 intonation units).¹ Based on this evidence, there seems to be little reason to believe that Finnish makes use of free NPs to a greater extent than other languages.

At present, however, we do not know enough about free NPs and the functions they serve in different languages to set up any typological claims about free NPs. Furthermore, in the studies we have so far, free NPs are sometimes discussed only in passing, as one type of syntactic construction (e.g. in Durie 1994 and forthcoming), or only one subtype of free NPs is being discussed with no information on other types of free NPs there may exist in the

1. Tao (1996) looked at NP intonation units, i.e. intonation units consisting of an NP only and found that in his Mandarin data, 28.7% of the intonation units were NP intonation units (Tao 1996:72). Croft (1997:12) reports that in his data from Wardaman, the percentage of NP intonation units was 21, but in his English data, it was only 14. It is important to note that Croft and Tao are counting differently from what I have done here: they have analyzed a given intonation unit as an NP intonation unit if it consisted of an NP, whereas I have looked at NPs that are syntactically independent. Tao notes (1996:80) that slightly more than half (53%) of the NPs in NP intonation units did not belong to a clause core but functioned as free NPs. From these figures we can estimate that over 15% of the Mandarin intonation units contained free NPs in the sense we have been using the term. In other words, the proportion of free NPs is higher in Mandarin than it is in my Finnish data (5.1%).

language (e.g. Herring 1989). This makes cross-linguistic comparison very difficult. It does seem clear that there are genre differences in the extent to which free NPs are used as a grammatical resource: for example, Durie (1994) found that free NPs were most frequent in his conversational data, but hardly occurred at all in a written folk tale also included in his data.²

In conclusion, we have evidence from a variety of languages for free NPs as a type of syntactic unit which is distinct from and not reducible to clauses. Although we do not have a clear crosslinguistic understanding of the different functions free NPs serve in languages, the two main functions proposed by Ono and Thompson (1994), namely referential and predicating, do account for the functions assigned to free NPs in the literature so far.

4.3 Case marking of free NPs

Finnish is a language with an elaborate case marking system, but not all of the 14 cases are used in free NPs. The nominative is by far the most frequent, whereas other cases are used quite rarely. Moreover, there are cases that are never used in free NPs.

Lambrecht (1981:78–84) discusses the case marking of topic and antitopic constructions in French, and notes that topics (preceding the predication) are not marked for case, whereas antitopics (following the predication) are, at least potentially (cf. Ashby 1988:210). Lambrecht (1981:80) and Ashby (1988:227) take this to suggest that antitopics (or “right-dislocations” in Ashby’s terminology) show a tighter syntactic bond to the predication. In Finnish, however, if the free NP is case marked with a case other than the nominative (the unmarked case), it does not mean that the free NP would be more tightly connected with a clause than are nominative free NPs. As I will show shortly, case marking may even add to the syntactico-semantic independence of the free NP.

2. Durie (1994) does not give exact numbers for free NPs in the different sets of data he has analyzed. He has data from a written folk tale, an oral narrative and a conversation; these three contained 128 free NPs in all (Table 1, p. 505). In the written folk tale there was one “trackable” free NP (Table 12, p. 515), whereas in the oral narrative there were 45 of them (Table 13, p. 515), and in the conversation there were 79 (Table 14, p. 516). Thus there remain 3 free “non-trackable” free NPs of which we do not know in which data set they occurred. Nevertheless, the comparison of trackable free NPs shows that there were great differences among the three sets of data.

The nominative is the default case for free NPs in Finnish. This is understandable for many reasons: as was discussed in Chapter 2, the nominative is unmarked in Finnish, i.e. there is no separate ending for the nominative. It is by far the most frequent case ending (see Table 2 above; see Du Bois 1987b for a discussion of the discourse use of zero-marking). It is no wonder then that it is used so frequently in free NPs also. The partitive is sometimes used for mass nouns (ex. 41) or for an unspecified (unbounded) quantity of countable nouns (see Section 2.4.3 for a discussion of the partitive).

- (41) Alko
- 1 Mauri: $\wedge$ *Samelin on hiili-agentti*.
 Samelin is coal-agent
 Samelin is a coal agent.
- 2 Raimo: ... *mm*.
 Uhhuh.
- 3 Mauri: [₁ *hii₁*]*l-i-ä* *ja* $\wedge$ *koksi-a*. ←
 coal-PL-PTV and coke-PTV
 Coal and coke.
- 4 (X): [₁(*THROAT*)₁]
- 5 Mauri: *lö-i* *mu-lle* *lapu-n* *käte-en* *ja*,
 hit-PAST+3SG 1SG-ALL note-ACC hand-ILL and
 He slapped a note into my hand and
- 6 $\wedge$ *punti-n* *tikku-j-a* *tasku-u* *ja* *sano et*,
 packet-ACC match-PL-PTV pocket-ILL and said that
 a packet of matches into my pocket and said that
- 7 *jos* $\wedge$ *hiil-i-ä* *tarvii-t* *nii*,
 if coal-PL-PTV need-2SG so
 if you need any coal,
- 8 $\wedge$ *aina* *halva-lla*,
 always cheap-ADE
 always cheap,
- 9 $\wedge$ *halv-emma-lla* *ku* *muu-t*.
 cheap-COMP-ADE than other-PL
 cheaper than the others.

In example (41) line 3, the free NPs *hiiliä* ‘coal’ and *koksia* ‘coke’ are mass nouns. In line 1, we have a predicate nominal clause. The interpretation of line 1 is that Samelin sells coal, and line 3 gives more specific information: not

just coal, but also coke. (Note that the recording was made in 1956 when coal was still widely used for heating in Finland). The free NPs refer to the items he deals in, namely coal and coke. The predicate nominal clause in line 1 serves to characterize Samelin as a coal agent (dealer), but it is important to note that there is no mention of the process of dealing, and no clause that the NPs in line 3 would belong to.

In the analysis of example (41), some might want to argue that there is no need to mention the process, because it is self-evident in the context. However, the NPs in line 3 cannot be analyzed as part of an elliptical clause, since there is no single predicate that could be claimed to be missing: in principle, the NPs could function as objects of the verb *myydä* ‘sell’, e.g. *Se myy hiiliä ja koksia* ‘He sells coal and coke’, but they could also function as parts of a possessive construction *Sillä on hiiliä ja koksia* ‘He has coal and coke’.

The oblique cases can also be used in free NPs to indicate orientation (ex. 42). The oblique case ending can be attached not only to nouns, but also to more adverb-like roots which only take a subset of the oblique case suffixes, and which do not inflect in the grammatical cases. Free NPs that are in an oblique case can sometimes form a construction together with a NP in the nominative. I will call this construction “theme and orientation”, where theme refers to the nominative NP, and orientation to the oblique NP (see Section 4.4.2 for further discussion). Consider the following example:

(42) Alko

1 Olli: *mi-tä-s sä maksa-t ^vuokra-a.*
 what-PTV-CLT 2SG pay-2SG rent-PTV
 How much do you pay for rent?

2 Aarre: *^kaheksan-tuhat-ta.*
 eight-thousand-PTV
 Eight thousand.

3 Olli: *ei lämmitys-kusta- —*
 NEG+3SG heating-expense
 No heating expenses- —

4 *siis vielä lämpö ^pää-lle.* ←
 so still heat top-ALL
 that means heating excluded (lit. heating to top (of rent)).

5 Aimo: *ei= ku ^kaikk-i-a-an.*
 NEG+3SG as all-PL-PTV-PX
 No but (eight thousand) in all.

- 6 Olli: <P *joo no se o* P>,
 yeah well 3sg be+3sg
 Yeah well that is —
- 7 Tor: *entä* [^]*sähkö*? ←
 what.about electricity
 What about electricity?
- 8 Aimo: <P *tietysti* [^]*sähkö eri-kse-en* P>. ←
 of.course electricity separate-TRA-PX
 Electricity excluded of course.
- 9 Olli: [^]*tietysti*.
 of.course
 Of course.

In example (42), line 4, the nominative NP *lämpö* ‘heating’ is the theme and the adessive adverb *päälle* ‘on top’ indicates its orientation. The speaker Olli uses this construction as a request for confirmation. In line 8, Aarre uses a similar construction, but this time the construction functions as a reply to Tor’s question in line 7 (which is also made in the form of a free NP, but it is a topic rather than a theme and orientation construction). In the construction in line 8, the nominative NP *sähkö* ‘electricity’ is the theme and the translative adverb *erikseen* ‘aside’ indicates the orientation. Note that the NP in line 2, *kaheksantuhatta* ‘eight thousand’, is an elliptical response to Olli’s question in line 1. Because it can be analyzed as a part of an elliptical clause, it was not considered to be a free NP.

Of the fourteen cases in Finnish, the genitive and the accusative are hardly ever used in free NPs.³ This is quite understandable in light of the syntactic and semantic characteristics of these two cases: they depend on some other element for their interpretation. Syntactically, they cannot occur alone, and semantically, they cannot be interpreted alone: the accusative is dependent on the verb it is an argument of, whereas the genitive is dependent on the head of the phrase.

3. In my data, the most infrequent cases like the comitative and the instructive do not appear at all in free NPs, but this is not a special characteristic of free NPs, because these cases do not appear anywhere in the data (with the exception of some lexicalized adverbs).

4.4 Discourse functions of free NPs

In this section, I will present a brief overview of the different discourse functions of free NPs. I distinguish between four different types of free NPs: (i) identifying, (ii) classifying, (iii) theme and orientation constructions, and (iv) topic constructions. I will discuss identifying and classifying free NPs together, because they are very close to each other. As was discussed above, Ono and Thompson (1994) divide free NPs into two groups, referential and predicating. Our types (i)–(iii) are all doing predicating work, whereas the last group contains free NPs that function to manipulate referents. Thus, Ono and Thompson's analysis of the functions of free NPs in their English data is valid for Finnish also. While accepting this broad distinction, I have done a more fine-grained analysis.

4.4.1 Identifying and classifying free NPs

Free NPs can serve to identify a referent, a proposition or a longer sequence of text. Identification, of course, comes very close to classification, since both establish a similar kind of relation between two entities. The two types differ in that classifying NPs name a class to which an entity belongs, whereas identifying NPs serve to equate two entities. I have also coded as a subtype of identifying free NPs those that label an entity, thereby identifying it. This type differs from the others in that it does not establish a relationship between two entities, but rather, identifies an entity by naming it.

Example (43) contains a classifying free NP. Before this sequence in the conversation, Tintti has said that she has bought two bottles of liqueur for Christmas, although it is still three weeks to Christmas. Leena warns her against the temptation to drink them before it is time, but admits that one can always get new ones from the store.

(43) Mother and Daughter

- 1 Tintti: *e-n mä tai-da viitsi-ä ava-[_1ta._1]*
 NEG-1SG 1SG be.likely-INF care-INF open-INF
 I don't think I will open them.

- 2 Leena: [_1 ai_1] tää,
 oh this,
 Oh you mean,

- 3 ⟨@ aio-*t* @⟩,
 aim-2SG
 you are going to
- 4 ...(0.3) esiinty-ä ^kieltäymykse-n kuningatta [₂re-na.₂]
 play-INF abstinence-GEN queen-ESS
 play the Queen of Abstinence. (Or: appear as the Queen of
 Abstinence)
- 5 Tintti: [₂@ @₂]@
- 6 ^*nii*. ...(0.4) [₃justiinsa.₃]
 so right
 Yes. Right.
- 7 Leena: [₃@ @ @₃]@
- 8 ^*luonne-rooli*. ←
 character-role
 A character role.
- 9 *no* ⟨@ *niin* @⟩ [₄@@₄]
 oh so
 OK.
- 10 Tintti: [₄@@₄]
- 11 ⟨@ *rehellinen* @⟩ *luonne-rooli*. ←
 honest character-role
 A pure (honest) character role.

In lines 3 and 4, Leena comments on her mother's plan to save the alcohol for Christmas, suggesting that by doing so she is playing "The Queen of Abstinence". The mother accepts this by starting to laugh (line 5) before Leena gets to the end of her utterance. She adds two response tokens (line 6) that serve to close the sequence, but Leena starts to laugh, adding a characterizing free NP (line 8). It is not easy to pin down what exactly is being characterized. It could be either the role of "Queen of Abstinence" or the mother's playing it, i.e. the whole proposition (lines 3 and 4). Interestingly enough, the role has not been mentioned as a referential expression, but rather as a characterizing phrase in the essive case *kieltäymyksen kuningattarena* 'as the Queen of Abstinence'. By using a free NP in the characterization instead of a predicate nominal clause, the speaker avoids linking the characterization to a referential expression, such as the copular subject. In line 11, Tintti accepts the characterization by repeating it with a modification: *rehellinen luonnerooli* 'a pure character role'.

The following example contains several free NP constructions that serve identifying and classifying functions (types i and ii). The example illustrates how identification of referents is being negotiated between the co-participants. The example comes from the Birthday party conversation, where Veijo, the only male participant in the conversation, is telling a story about somebody he had met. Veijo introduces him as an athlete but doubts whether the other participants know of him. This is a challenge for the others that invites them to show off their knowledge about sports. (Only those lines that will be focused on have a morphological gloss.)

(44) Birthday

- 1 Veijo: *<X tapahtunu X> kerran kans ku <X tuli X> iso,*
(It has) happened once (to me) that there came a big
- 2 *iso <X juutas X> vastaan ^Lappeenrannan kadulla illalla.*
big guy towards (me) on a street in Lappeenranta (a little town in Finland) at night.
- 3 *mutta ette te ku- kukaan tunne sitä*
but you won't know him
- 4 *[₁ kuitenkaan, mun on ^ihan turha kertoa. ₁]*
anyway, so it's useless that I tell you (about him/his name).
- 5 Sanna: *[₁ se oli Saimaa-n Pallo-n, ei ku mikä se nyt oli ₁].*
it was Saimaa-GEN Ball-GEN no but what it now was
He was from Saimaan Pallo (a local ice-hockey team in Lappeenranta), no what was it now.
- 6 Veijo: *ei ku se oli ^urheilija niin te e-tte ^kukaan*
no but it was athlete so 2PL NEG-2PL nobody
No, he was an athlete so you wouldn't
- 7 *[₂ tie-dä <X kuitenkaan X> ₂],*
know-INF anyway.
know anyway.
- 8 Leena: *[₂ no kuka ₂],* ←
oh who
Well who?
- 9 Veijo: *kuka se on.*
who he is.
- 10 Mella: *no kerro,*
Just tell (us),

- 11 *mä voim* [₃ ^tietää₃].
I may know (him).
- 12 Mella: [₃ *joku* ^*Sai*₃] *maa-n Pallo-n*, ←
some Saimaa-GEN Pallo-GEN
Some (guy) from Saimaan Pallo,
- 13 Veijo: *Steve* [₄ *Ovett*₄]. ←
Steve Ovett.
- 14 Raita: [₄ *ai* *tää* *SaiPa*₄]. ←
oh this SaiPa
Oh (you mean) this SaiPa (abbreviation for Saimaan Pallo).
- 15 Leena: <*HI* ^*kuka HI*>? ←
Who?
- 16 Veijo: ... *Steve Ovett*. ←
Steve Ovett.
- 17 Leena: ... ^*juoksija*. ←
Runner.
- 18 Veijo: *no nii*,
All right.
- 19 Leena: [₁ <*F* *no* [₂ ^*nii*=*F*] _{1,2}].
All right.
- 20 (X): [₁ <*X* *X*> ₁] *näi*=*näytet* [₃ *tii*₃].
so we showed (him?).
- 21 Sanna: [₂ *hyvä* ^*Leena*]₂]
Great Leena!
- 22 Leena: [₃ *no n=ii*₃].
OK. (There we are.)
- 23 Noora: *mä* ^*oisin veikannu* ^*jääkiekkoilijaa*.
I would have guessed an ice hockey player.
- 24 Veijo: *ehe se on maailma-n* [₄ *kuuluis-imp-i-a*₄],
no 3SG is world-GEN famous-SUP-PL-PTV
Nope. He is one of the world's most famous
- 25 Leena: [₄ <*F* ^*maili F*>₄].
mile
The mile!

- 26 Veijo: *mm- mailere-i-[₅ta ja=₅].*
 miler-PL-PTV and
 mm- milers and.
- 27 Leena: [₅⟨HI ^mm. HI⟩₅]
 Mmm.

The example starts out with a lengthy sequence where the participants discuss the identification of the person that Veijo has met. In line 5, Sanna suggests a characterization with a predicate nominal clause. However, the NP containing the predicate nominal is not complete since the head is missing and there is only a genitive modifier *Saimaan Pallon* ‘Saimaan Pallo-GEN’. Veijo’s line 6 contains another predicate nominal clause which characterizes the referent as an athlete. In his utterance Veijo ignores Sanna’s candidate understanding of the referent as somebody from the local sports team (line 5) by saying that because he was an athlete the co-participants won’t know him. Leena requests an exact identification with an utterance containing the question word *kuka* ‘who’ (line 8). Mella’s line 12 repeats the characterization given already by Sanna in line 5. Again the NP only contains modifiers projecting a head noun (such as *urheilija* ‘athlete’ or *jääkiekkoilija* ‘ice hockey player’), which is lacking. Veijo identifies the person in line 13 by naming him. Up to this point in the conversation, the NP intonation units (lines 8, 12 and 13) have been designed so that they fit the predicate nominal clause construction given first by Sanna in line 5 and then by Veijo in line 6. We could say then that the NP intonation units contain elliptical clauses. They are all concerned with the characterization and identification of the person Veijo had met. However, Raita’s line 14 is different: she checks her understanding of the referent of the genitive modifier *Saimaan Pallon* in Mella’s utterance (line 12) by giving a paraphrase: *SaiPa* is the abbreviation used for the team in sports newscasts. Raita’s NP *tää SaiPa* ‘this SaiPa’ picks a referent, and with the turn-initial particle *ai* she construes her utterance as a confirmation check (cf. Sorjonen 1989:131): she checks whether Saimaan Pallo can be equated with SaiPa.

In line 15, Leena repeats her request for identification. The request is made with high pitch (marked with ⟨HI⟩ in the transcript), compared to Leena’s normal pitch range. The use of the high pitch functions to display that she has not heard Veijo’s earlier reply (line 13). Veijo repeats the name (line 16), thereby identifying the person. In her response (line 17), Leena shows that she has understood the reference as she names the category the referent belongs to (*juoksija* ‘runner’). In other words, with the use of the proper name Veijo picks

a specific referent, and consequently, the co-participant displays the understanding she currently has of the intended referent: “you mean Steve Ovett, the runner”. Veijo accepts Leena’s understanding with his response *no niin* ‘alright’ (line 18). It is an interesting question whether the NP *juoksija* ‘runner’ is general or particular. One could think that it relates the referent just mentioned, Steve Ovett, to the class of runners. But one could also argue that it identifies the referent as the runner that we know under that name. In my view, these two interpretations are both available.

Leena’s ability to identify the referent is greeted with joy (lines 20–22). In line 23, Noora admits that she had searched for the intended referent, Steve Ovett, in the wrong category, among ice-hockey players. Veijo corrects her (line 24) by characterizing the referent with a predicate nominal clause, but in the middle of the characterizing phrase, Leena breaks in with a free NP (line 25). The exact point of Leena’s intrusion is interesting because it is not a transition relevance place (Sacks et al. 1974): Veijo has just uttered a genitive modifier *maailman* ‘the world’s’ which projects more. The projection is quite specific, namely a superlative adjective and a head. Thus, Leena’s NP is syntactically not a possible continuation of the NP Veijo has started to construct. Furthermore, it is prosodically produced so that it could not be interpreted as a co-construction. Namely, Leena is using a loud voice (marked with ⟨F⟩ in the transcript) and relatively high pitch which is in contrast with Veijo’s rather soft voice. Lerner (1991, 1996), who has studied co-constructions in English conversations, has found that they typically occur in what he calls “compound turn-constructional units” consisting of a preliminary component, such as an if-clause, and a final component, e.g. a then-clause. The NP that Veijo has started to produce in line 24 has a complex structure, but it is not structured as a compound turn-constructional unit with two parts that the participants could co-construct.

Leena’s line 25 is produced in overlap with the modifying adjective in Veijo’s utterance. Although Veijo is speaking himself, he does show that he has attended to Leena’s NP, as there is a slight hesitation before Veijo comes to the head of the phrase *mailereita* ‘milers’ (line 26).

Example (44) illustrates nicely how the identification of a referent is negotiated in discourse: during the whole sequence, the co-participants are occupied with the identification of the person Veijo has met. For example in lines 15–16 we have a question-answer pair, which is followed by a candidate understanding of the answer (line 17) and a confirmation (line 18). The NP in line 17 serves to pin down the referent. The NP in line 25 (*maili* ‘the mile’)

identifies the field of sports with which the person is occupied, thus adding to the co-participants' understanding who Steve Oveti actually is.

In conclusion, free NPs are often used to classify or identify something which may have been referred to in the prior text with a NP or may be a larger sequence like a proposition or a stretch of prior text. In the flow of discourse, they provide the co-participants an opportunity to check their understanding of the reference intended by the primary speaker, or to evaluate a prior sequence. Thus, they are backward-looking, as opposed to topic constructions (Section 4.4.3), which are forward-looking (Ono and Thompson 1994). Although identifying and classifying NPs do predicating work in discourse, they are oriented towards negotiating reference rather than making predications that would initiate a new activity.

4.4.2 Theme and orientation

As was noted in Section 4.2, locative NPs can sometimes form a construction together with a NP in the nominative. I have called this construction “theme and orientation” (type (iii) in our typology, see earlier this Section 4.4), where the theme is the nominative NP, and the orientation is the locative NP. The whole construction serves to make a predication about the theme. Theme and orientation constructions differ from the other predicating free NPs discussed above (identifying and classifying) in that they start something new rather than just look back to prior talk.

In the theme and orientation construction the local case ending expresses either a location or a direction towards or away from something. In a way it functions like a predicate which has its own arguments, namely the stem of the noun to which the case ending is attached and, optionally, a theme (Helasvuo 1991). In Finnish linguistics, Siro (1964) has proposed that case-endings are quasi-predicates which take arguments. In an intransitive clause, the case-ending takes the object as its argument (“refers to”, in Siro’s terminology), and in a transitive clause it takes the subject. Consider the following examples (the morphological gloss is not exhaustive, but only highlights the relevant points).

- (45) a. Siro (1964:28)
Asiakas astui toimisto-on.
 client stepped office-ILL
 The client stepped into the office.

- b. Siro (1964:28)
Sisaret veivät hänet kylä-än.
 sisters took her village-ILL
 The sisters took her to the village.

In (45a) it is the subject referent (client) which moves to the location indicated by the illative-marked constituent. The clause is an example of an intransitive clause, and accordingly, the illative marker takes the subject as its argument. In (45b), it is the referent of the object *hänet* ‘her’ that moves. The clause is transitive, and the case ending takes the object as its argument.

Theme and orientation constructions make use of this meaning potential that the case endings have. Consider example (46).

(46) Alko

- 1 Aarre: *jos mä otan lauantai-iltana ^viinaa.*
 If I drink booze on Saturday night.
- 2 *ja mä otan oikeen ^reilusti,*
 and I drink a lot,
- 3 *että mä nyt tunnen sen [₁ oi₁]kein päässä ni,*
 so that I feel it really in my head,
- 4 Olli: [₁ ^nii.₁]
 Right.
- 5 Olli: ^sika-jurri-in [₂ nii.₂] ←
 pig-drunk-ILL so
 (To become) drunk as a pig, right.
- 6 Aarre: [₂ ^nii.₂]
 Right.

In example (45) line 5, the free NP *sikajurriin* ‘pig-drunk’ is in the illative case, which generally expresses movement towards something. Here the movement can be understood as movement towards the state of being “blind drunk” (note that *jurri* is a noun referring to the state of being drunk). The free NP serves to make a predication about Aarre and his drinking, and thus, Aarre is the assumed theme, but it is not expressed. It is important to note that there is no clause in the context that the free NP could belong to, but rather, it forms a construction all by itself.

Constructions with a theme and an orientation are quite common in the written language also, but they mainly occur in very specific contexts, namely in titles and headings. Consider the following examples.

- (46) Newsweek 11/3/97
The passport to your future
- (47) Thomason and Kaufman 1988:325
On orderliness or the lack of it in the rates of linguistic change in English

In example (46), passport is the theme and to your future is functioning as an orientation, whereas in example (47), there is no overt theme but only an orientation.

In their paper on free NPs in English, Ono and Thompson (1994) have an example which contains a theme and orientation construction:

- (48) Ono & Thompson (1994): example 11
- | | | |
|------|--|-------|
| 1 A: | @@@ @so @she wants to [go out on the balcony], | |
| 2 L: | | [(H)] |
| 3 A: | I grab her again, | |
| 4 | [and I go], | |
| 5 L: | [(Hx)] | |
| 6 A: | (H) no kids on the balcony. | ← |

In example (48) line 6, the NP *no kids* functions as the theme and the prepositional phrase *on the balcony* predicates its orientation. Ono and Thompson note (1994: footnote 3) that some people have objected to this example saying that it is grammaticized as a kind of prohibitive quote. Ono and Thompson agree with this, but continue, “[W]e take its grammaticized use as a prohibition as supporting our claim that predicating NPs have a life of their own apart from any clause structure.” As we have seen, constructions with a theme and orientation are not restricted to specific contexts, such as prohibitive quotes or signs, but can appear in various contexts.

It is interesting that case marking, which in itself is one of the defining characteristics of nouns morphologically, can be used to bring predicating power (e.g. motion towards something) to the construction. The locative case marker (or the preposition) can be analyzed as functioning as a predicate which takes the stem of the noun which it is attached to as its argument and, optionally, a theme. It is important to note that identifying and classifying NPs, although they also do predicating work in discourse, are primarily used to negotiate reference or to characterize already established referents, whereas constructions with a theme and orientation can be used to make predications not concerned with negotiating reference, but instead, initiate something new.

4.4.3 Topic constructions

So far, we have discussed predicating free NPs. In this section, I will look at free NPs that do referential work in discourse. In the literature on free NPs, those free NPs which serve to present a topic into discourse have been quite widely discussed (Chafe 1976, Keenan and Schieffelin 1976, Lambrecht 1981, 1988, Geluykens 1987, 1988, 1992, Ashby 1988, Tao 1992, Ono and Thompson 1994). Topic NPs are referential, and in terms of the flow of discourse, topics are forward-looking, i.e. they are interpreted in the light of what follows (Ono and Thompson 1994). At the same time, they project more talk on that topic.

Chafe (1976) defines a topic construction as a construction in which a topic is first mentioned in a NP, and then in the following clause there is a predication about it containing an anaphoric pronoun that refers to the topic NP. In my work, the term topic is understood in a slightly looser way: topic NPs need not be anaphorically bound to an element in the immediately following or preceding clause. A distinction is usually made between topics (preceding the predication) and antitopics (following the predication; see e.g. Lambrecht 1981). “Antitopic” may be a somewhat unfortunate term in that it seems to imply that it works in a way contrary to topics. However, both function to organize larger stretches of discourse, and they refer to referents that usually continue to be talked about. In this section, I will discuss topic constructions with respect to their discourse functions. I will not maintain a distinction between “topics” and “antitopics”.

Geluykens (1988, 1992) discusses topic constructions in English in the larger framework of negotiating reference. He does not, however, use either the term topic or free NP, but instead, speaks about dislocations. This terminological choice can be seen as reflecting his analysis of topic constructions as clausal constituents that are dislocated from where they “belong”. I consider this line of analysis quite problematic: given the linear organization of natural speech, how can something be dislocated from something that has not yet been uttered? The analysis of topic constructions as dislocations reflects a more general approach that views spoken discourse as a finished product rather than an on-going process.

According to Geluykens (1988, 1992) topic constructions are used to negotiate reference. Geluykens (1988:2) proposes that topic constructions involve a process with three stages:

- i. a referent is introduced with a free NP (the “left-dislocation”),
- ii. the referent is then acknowledged by the co-participant(s) (“hearers” in Geluykens’ terminology),
- iii. the speaker proceeds to make a predication about the introduced referent.

Thus, according to Geluykens, the main function of topic constructions is to introduce referents. According to him, topic constructions introduce referents that are new (or irrecoverable, to use Geluykens’ terminology) from the previous discourse. This finding is supported by Herring’s (1989) studies on topic constructions in Tamil. Herring (1989) found in her study on topic constructions (“verbless presentations”) in Tamil that they are specialized for presenting new information in Tamil: almost 90% of them introduce new referents.

In contrast, Ashby (1988) found in his study on topic constructions in French that only half of them refer to new referents (1988:212). According to him, introduction of new referents is not the primary function of topic constructions in French; instead, they are primarily a topic-shifting or topic-creating device (Ashby 1988:217), i.e., they function to structure larger stretches of discourse.

Both Ashby (1988) and Geluykens (1992) found that the referents of topic constructions usually continue to be talked about in the subsequent discourse; however sometimes, as Geluykens (1992:155) notes, the referent is not continued as such, but serves to bring in some other referents that are linked to the referent of the topic construction (cf. Chafe 1976:50).

Interestingly enough, none of my examples manifests what Geluykens identifies as the prototypical three stage referent negotiation pattern for topic constructions. Of course, my database of topic constructions is not very large (only 12 NPs), and more data may reveal different patterns. In some cases, topic constructions served to pick a specific referent from a set that had been mentioned before. In other words, the topic NP serves to focus on a specific referent that may be inferable from prior text. Thus, topic constructions seem to be functioning in a different way in Finnish than in the English data Geluykens has collected. Consider example (49), which comes from the conversation between two brothers.

(49) Brothers

- 1 Pekka: *ku me hiihdettii ni, ... meidä ohi meni semmonen lumikelkka,*
when we were skiing, there was a snowmobile that drove past us
- 2 *mi-s ol-i kaks ukko-o.*
REL-INE be-PAST+3SG two man-PTV
where there were two men.
- 3 Simo: *nii,*
Yes.
- 4 Pekka: *... no sit alko tulemaa isoi mäkii eteen ja,*
... Well then we got into a hilly terrain with lots of big hills and,
(lit. “there started to come big hills towards us”)
- 5 *... ja sit me alettii tekee lumiluolaa semmoseen kinokseen ja,*
... and then we started to make a cave into the snow and,
- 6 *... se oli perhanan kovaa lunta,*
... It was damn hard snow.
- 7 *ei me saatu lapiol mitää, mitää reikää siihen jäätyneeseen lumeen*
we couldn’t make any, any hole in that frozen snow with a spade
- 8 *ja, ... sit sielt, ... tuliki takasi se lumikelkka sieltä,*
and, ... then there... came back the snowmobile from there,
- 9 *ja se toinen, ... kaveri siin kyydi-ssä,* ←
and 3SG other guy there lift-INE
and the other... guy there in the snowmobile,
- 10 *se kysy et, jos me ote-ta-an lifti*
3SG ask-PAST+3SG that if 1PL take-PSS-PERS lift
he asked if we would like a ride
- 11 *semmo-see, ... hyti-lle.*
a.kind.of-ILL cabin-ALL
to a kind of a cabin.

In the example lines 1 and 2, Pekka introduces two men who had driven past them with a snowmobile while they were skiing. Simo acknowledges the introduction with the particle *nii*, roughly translated as ‘yes’ (line 3). The particle treats the prior talk as still incomplete and shows an expectation that the talk should be continued (Sorjonen 1997:436). After a short description of their skiing, Pekka returns back to the snowmobile (line 8), and introduces the other one of the two men on the snowmobile with a topic NP *se toinen kaveri* ‘the other guy’. The referent is referred to again in line 10 with the anaphoric 3rd

person pronoun *se* in the beginning of the clause, now functioning as the transitive subject. Thus, the topic construction serves to focus on a referent that has been mentioned in the prior discourse.

A topic construction may also follow the predication concerning it (in the literature, this is usually termed an “antitopic”). Consider the following example:

(50) Birthday

- 1 Leena: *sitte mu-lla oli ollu se viiko-n ja*,
 then 1SG-ADE had had 3SG week-ACC and
 then I had had it (a record) for a week and
- 2 *ja sitte se oli ollu = tä-stä viiko-sta nin*,
 and then 3SG had been this-ELA week-ELA so
- 3 *viis päivä-ä laina-ssa Erja Honkasalo-lla*.
 five day-PTV loan-INE Erja Honkasalo-ADE
 and I had lent it to Erja Honkasalo for five days out of this
 oneweek (lit. it had been out of this one week five days on loan
 to Erja Honkasalo).
- 4 (X): @@
- 5 Leena: *se ^levy ja*, ←
 3SG record and
 the record and,
- 6 Raita: *^nime-t*, ←
 name-PL
 Names!
- 7 Mella: @@@

In example (50), line 5, there is a free NP *se levy* ‘the record’ which could be termed an “antitopic” construction. In lines 1 and 2, there is a co-referential pronoun *se* ‘it’. Geluykens (1987) discusses constructions like the one in example (50) as tails. He analyzes them in terms of a repair mechanism which allows the speaker to identify a referent. The NP in line 5 can be understood in terms of a repair; however, it does not seem to serve to identify a referent as much as to disambiguate the reference. The free NP is co-referential with the subject of the preceding clause, which is the 3rd person pronoun *se* ‘it’ referring to a non-human referent. As we saw in Chapters 2 and 3, human referents are quite pervasive in the subject role. There is also evidence in the literature (see especially Givón 1983) that human referents have higher persistence in discourse, i.e. they continue to be talked about. In the clause in lines 2–3, there is a NP *Erja*

Honkasalolla referring to a human referent, but not in the subject role; instead, it is in the oblique role. In reference to the loan, the speaker uses a construction that leaves the loan-giver, i.e. the speaker herself, unspecified. Thus, there are two human referents in the context, but neither of them serves as the subject (note that the speaker does refer to herself in the preceding clause in line 1, but again, not in the subject role, but with a locative NP *mulla* ‘on me’). I would like to suggest that the free NP in line 5 is needed to disambiguate reference because there are these two human referents that are much more likely to be continued as discussion topics than the non-human referent.

According to Geluykens (1987:126), the “tail” constructions in his data were neither given nor new, but inferrable from previous discourse. However, the referent of the “tail” construction in example (50) is clearly given: it has been talked about for some time and has been mentioned several times.

In example (50) line 6, there is another free NP *nimet* ‘names’ which is interesting in terms of topic shift: it relates to a topic that had been discussed a while ago, namely the fact that the speakers should not use names of people because they are being taped. Thus, the free NP does metatextual organizing work: it points to a topic in prior discourse and, at the same time, it functions to make the participants aware of the setting of the conversation. The free NP in line 6 serves to organize the flow of discourse just like any other topic construction. It also illustrates the fact that topic NPs need not be in an anaphoric relationship to a clausal argument in the immediate context.

In conclusion, the topic constructions in my data serve to focus on a specific referent which may already have been mentioned in the discourse or which may be member of a set of referents that has been mentioned. In comparison to the vast majority of NPs that are clausal arguments, the syntactic form of topic NPs as NPs that are unattached to clausal argument structure suits well the function of focusing on or highlighting a referent.

4.4.4 Other

Sometimes an NP may be like a free NP in that it is not an argument of any verb, but it may not conform to the regular patterns of free NPs described above. Rather, it may be a result of a change of plan or an abandoned structure. Example (51) serves to illustrate this.

(51) Birthday

- 1 Veijo: *ei-päs, se [oli väärä-s],*
 NEG-CLT 3SG was wrong-INE
 No, no, she was wrong.
- 2 Raita: [*kai se nyt*],
 perhaps 3SG now
 Perhaps she/he/it now —
- 3 Veijo: *ei-hä tää nii ollu ku,*
 NEG-CLT this so been but
 It wasn't like that but

In example (51) line 2, the NP *se* 'she/he/it' is part of a structure that is never completed. It is overlapped by Veijo's utterance (line 1).

If free NPs are thought of in terms of schemas (see above Section 1.1), cases like example (51) line 2, do not represent a free NP schema, and thus, are not really free NPs, but rather, some other clausal schema that is only partially realized.

4.5 Summary

In this chapter, we have seen evidence for the fact that speakers can use NPs as a resource to form syntactic units that are unattached from any clausal argument structure. We also saw that there is evidence for free NPs in other languages as well. However, it seems that the research on free NPs is still too sparse to allow for any cross-linguistic generalizations.

Free NPs can be used to perform a variety of functions, which range from typical noun functions relating to the manipulation of referents to less prototypical ones such as orientation and classification. In Section 4.4.3, I showed that referential free NPs can be used to help to highlight a referent or to focus on a referent that is a member of a larger set of referents already under discussion. We could say that topic NPs serve functions that are closest to the prototypical functions of nouns as identified by Hopper and Thompson (1984), namely, manipulating reference. Among the predicating free NPs, identifying and classifying free NPs (Section 4.4.1) function either to characterize referents or to disambiguate the intended referent(s). In other words, they serve mainly in negotiating reference, whereas constructions with a theme and orientation (Section 4.4.2) make predications that may initiate something new, rather than look back to check the understanding of prior talk.

CHAPTER 5

Intonation and syntactic structuring

As was discussed in Section 1.4, the data for this study have been transcribed into intonation units, identified on the basis of several prosodic cues. Intonation units are perceptual units, defined as a stretch of speech uttered under a single coherent intonation contour (Chafe 1987, Du Bois et al. 1993). In addition to this rather broad characterization, there are some further cues to help in the identification of intonation units, such as pitch reset in the beginning of the unit, acceleration in tempo on the initial unstressed syllables, prosodic lengthening of the final syllables, and pausing (Du Bois et al. 1993; see also Cruttenden 1986). In this chapter, I will look at how these perceptually identified units are structured syntactically.

Many studies on intonation make the assumption that intonational grouping reflects syntactic structure directly. In intonation studies that focus on single sentences (see e.g. Pierrehumbert 1980, Ladd 1986), the question of the relationship between intonation and syntax does not even arise as the scope of intonational patterning has been limited a priori to the sentence. Thus, the assumption that there is a direct relationship between syntax and intonation can be left implicit, whereas in studies that are more oriented to larger stretches of discourse, it may be stated explicitly (see e.g. Halliday 1985, Cruttenden 1986). For example, Cruttenden (1986:130) states: “[I]ntonation-groups generally correlate with major syntactic constituents, although a good deal of choice is available to speakers concerning which constituents intonation-groups should correspond with.” By “major constituent boundaries” Cruttenden (1986:37) means boundaries between clauses and between the subject and the predicate.

Thus, according to Cruttenden, intonation unit boundaries should occur either at clause boundaries or between the subject and the predicate. We will return to this point below (Section 5.2). In many studies it is assumed that in an unmarked case the intonation unit is coextensive with one clause (see e.g. Halliday 1985:274). In this chapter, I will consider whether these claims can be supported by analysis of my Finnish data.

As was discussed in Chapter 2 (especially Section 2.1), the clause as an

analytic category is often assumed to have a shared, uncontroversial meaning that needs no discussion.¹ I tried to show that syntactic relations are clearest in the clause core between the predicate and its core arguments. In this chapter, I will further show that the clause core emerges as a unit also in terms of intonation patterns. In Chapter 4, I showed that the noun phrase is very robust as a syntactic category: not only may it function as a clausal constituent, but also as a free construction, the so-called free NP (or unattached NP, see Chapter 4). As we will see shortly, NPs show strong intonational unity as they are almost always produced in one intonation unit.

In this Chapter, my aim is to show that intonation is an important means of combining elements together to form units in discourse, on a par with grammatical and semantic grouping. As we will see shortly, the grammatical and intonational groupings most often coincide. But we may ask why this is so. One possible explanation is that it is another manifestation of iconicity in grammatical organization. The iconicity principle has been stated in terms of several different correlations, such as “the linguistic distance between expressions corresponds to the conceptual distance between them” and “the temporal-physical distance between chunks of linguistically coded information correlates directly to the conceptual distance between them” (see e.g. Haiman 1983, 1985a, 1985b, Givón 1991). In intonational grouping, the temporal aspect is of course crucial, and it includes not only linear organization but also rhythm and pausing. As we have discussed, pauses and rhythmic changes typically occur at intonation unit boundaries. Given the iconic correlation between temporal distance and conceptual distance, we could hypothesize that intonation unit boundaries are very likely to co-occur with boundaries of grammatical constituents, such as NPs and the clause core.

In this chapter, my main research question has to do with what intonation can tell us about the emergence of clauses and NPs as syntactic units. I will try to answer this question by first comparing clausal constituent structure with

1. For example Croft (1995), in his study on the relationship between intonation units and grammatical structure, notes that clauses and phrases “are generally clearly identifiable in spoken language” (1995:841), and then proceeds — without ever defining what he means by “clause” — to present tables on the relationship between intonation units and grammatical units, such as the clause. From the tables we can read, *inter alia*, that 38.1% of the intonation units in his corpus consist of simple clauses with subject (this is the largest group), whereas only 1.8% consist of a clause + adjunct (Croft 1995:845). From this we may infer that the category of clause is taken to encompass only the clause core, although this is not explicitly discussed.

intonational grouping (Section 5.1). The claim that intonation unit boundaries should co-occur with boundaries for grammatical constituents (e.g. Cruttenden 1986) will be considered in the light of the Finnish data. Then we will look at the relationship between clauses and intonation units (Section 5.2) to see whether clauses indeed are presented in one intonation unit as has often been claimed (e.g. Halliday 1985).

5.1 Intonation and constituent structure

Corpus studies on English (Crystal 1969, Brown 1977, Croft 1995) indicate that the vast majority of intonation unit boundaries co-occur with boundaries for grammatical units. Whether or not these units are considered to be constituents depends on how constituents are defined. In the Finnish data, we can find strong support for the category of NP in discourse since NPs are rarely split into two intonation units. In this section, I will first discuss briefly how constituency could be understood in order to account for the patterns found in the discourse data. Then we will look at the results of the analysis of the Finnish data.

5.1.1 The emergence of constituent structure in discourse

In this section, I will outline a discourse perspective on constituency. The approach is based on Langacker's (1995, 1997) views on the emergence of constituency.

Langacker (1997) has proposed that constituency emerges from more basic phenomena, such as conceptual grouping, phonological grouping and symbolization. He argues (1997:3), "[C]onstituents are emergent in character, and while certain kinds of constituents are both natural and highly prevalent, they are not a *sine qua non* for grammatical viability." Langacker uses the notion of "classic constituency" (1995, 1997) to account for the most prototypical cases where the conceptual and phonological groupings converge. In a classical grammatical constituent, there is a valence-link between the component structures, i.e. they form a conceptual constituent, and the component structures are temporally contiguous and thus, form a phonological constituent (Langacker 1997:13). Langacker's model of constituency allows for different kinds of groupings; for example, in an ordinary transitive clause it is quite natural to group the verb with the subject, and not only with the object (Langacker 1997:8–9).

In my view, there are (at least) two aspects in Langacker's proposal that are relevant for the present work: First, the possibility of having different kinds of groupings among the elements allows us to account for patternings found in the data (see Section 5.2). Secondly, it is an important insight that some constituents may emerge more strongly than others, i.e. they may form classic constituents. As I have shown in Chapters 2 and 4, the category of NP is very robust in Finnish, and in this section, we will see that NPs show strong unity in terms of intonational grouping as well. I will also show that the Finnish data does not provide such strong support for the category VP in Finnish.

5.1.2 Intonation and constituency in Finnish

In this section, we will study the relationship between intonation and constituency on the basis of an analysis of the Finnish data. I will first present the overall pattern, and then I will discuss some exceptional cases.

In the Finnish data, 98.5% of intonation unit boundaries were also constituent boundaries. In counting the percentage, I did not assume the category of verb phrase, but instead, I just looked at boundaries of noun phrases and verbal groups (consisting of the predicate verb and possibly some infinitival forms). In comparison, Croft (1995) found that in his data from oral English narrative, 97% of intonation units were also grammatical units, and in Helasvuo (1997a) I show that in a database consisting of American-English multiparty conversations, in 98.8% of intonation units the boundary coincides with a constituent boundary. Thus, the results indicate that there is a strong tendency in spoken discourse for intonation unit boundaries to co-occur with boundaries of grammatical units. But even in the remaining cases where a constituent is split into two intonation units, we can find distinct patterns that I will illustrate below.

In most cases of constituent split in the Finnish data, the intonation unit boundary occurs after a determiner but before the head noun. In the data, there were 30 NPs that were split across two intonation units, and in 27 of these split NPs, the split occurred between the determiner and the head, and in the other 3 cases, it was between a genitive modifier and the head. Most often the determiner is the definite article *se* or the pronominal adjective *semmonen* which is often used to initiate a description of a referent and thus, to signal that there is more information coming (Erringer 1996: 77–78). It may seem odd that these determiners should be in a separate intonation unit from the head noun, but the pattern becomes more understandable if we consider the syntactic characteristics of noun phrases in Finnish.

As discussed in Chapter 4, in Finnish noun phrases, nearly all determiners and modifiers precede their heads and agree with them in case and number. Thus, the determiners and modifiers show the case and number of the whole construction, and in terms of syntactic processing they give an idea of how the clause as a whole will be structured, even though they may not contain enough semantic and discourse referential information to resolve the reference. This means that the form of the modifiers projects the case and the number of the whole construction, and also, gives hints about the syntactic role in which the construction functions (on projectability and constituent structure, see Himmelmann 1997). Consider example (52), which is an excerpt of a sequence of stories about people's first visits to their boy/girlfriend's parents.

(52) Birthday

- 1 Leena: *sit se o tietenki men-ny ^autta-maan tä-tä,*
 then 3SG had naturally go-PCP help-INF this-PTV
 then of course she had gone to help this,
- 2 ... (THROAT) *^toivottavasti tuleva-a anoppi-a*
 hopefully future-PTV mother.in.law-PTV
 hopefully future mother-in-law

In example (52), the determiner *tätä* 'this-PTV' appears in the same intonation unit as the verbal group *oli mennyt auttamaan* 'had gone to help'. It is a determiner that serves to mark the NP as identifiable. According to Laury (1997:62), *tämä* (partitive *tätä*) is typically used for referents which the speaker is currently focusing on and manipulating. In the example, the referent has not been mentioned before, but it is possible to treat it as identifiable by virtue of its sequential position in the series of stories about the first visits to the girl/boyfriend's parents. In this shared frame, the parents can be identifiable. In reference to the mother, the speaker uses a description that she has obviously made up on the spot: *tätä toivottavasti tulevaa anoppia* 'this hopefully to-be mother-in-law'. The "on-the-spot" quality of the descriptive phrase is reflected in the fact that it is split into two intonation units and that the speaker stops to clear her throat.

As we can see from example (52), the modifier *tätä* already shows the case and number of the head (partitive singular). In principle, it could form the object NP all by itself. However, it does not provide enough referential information. Also, the terminal intonation contour is continuing (marked with a comma) rather than final (marked with a period). Thus, both prosodically and grammatically, it projects more.

As Laury (1997) has shown, the definite article *se* has only recently (within the last hundred years) been grammaticized in Finnish, and on a par with the article usage, it is still used in its older function as a demonstrative pronoun. Thus, it can be used either as an independent demonstrative pronoun or as a definite article in a noun phrase. It is important to note that in the definite article usage *se* does not usually receive accent, but the head does. In the cases that I have analyzed as constituent splits the determiner was never accented, thus indicating that the two parts indeed formed a constituent even though the constituent was split into two intonation units. Consider example (53), which comes from the conversation between Tintti and Leena, who are mother and daughter. Before this sequence, there has been talk about spending money, and the daughter Leena reports on her friend telling about her mother's ways of dealing with money.

(53) Mother and daughter

- 1 Leena: *se oli aina tarkka rahoistaan*,
 She (Leena's friend's mother) was always careful with her
 money.
- 2 *mut sitte se vähän innostu nyt kahdeksankymmentluvulla ja*,
 but then she got a little excited now in the 1980's and,
- 3 *kuinka ^sääli oli että*,
 how pity was that
 What a pity it was that
- 4 ... (0.6) *että tota se ei=*,
 that er it NEG+3SG
 ... that er... she didn't have time to use
- 5 ...%% *si-tä=*, ←
 it-PTV
 the
- 6 ... (0.5) *min- ^minkki-turkki-a ehtiny käyttää ku*, ←
 mink-coat-PTV have.time use than
 ... min — mink coat more than
- 7 *pari ^talve-a*
 couple winter-PTV
 a couple of winters (before she died).

In example (53) we have an object NP *sitä minkkiturkkia* ‘the mink coat’ that is produced in two intonation units (lines 5 and 6). In line 4, there is the subject *se* ‘she’ and the negative verb *ei* which shows agreement with the subject (3rd person). They are followed by the object NP that consists of two parts, the determiner *sitä* (which is the partitive form of the definite article *se*) and the head noun *minkkiturkkia* ‘mink coat’. Between the determiner and the head, there is a pause of 0.5 seconds, and a slight hesitation. The determiner *sitä* shows the case of the head and number of the phrase, namely the partitive singular.

Interactionally the split constituent does very interesting work: Just prior to the lines we have been focussing on, Leena has told Tintti that the friend’s mother had been using money sparingly but that she had changed her attitude towards money just before she died (lines 1–2). In the Finnish culture, mink coats are status symbols for women, symbolizing a luxurious life. Thus, the mink coat referred to with the split constituent exemplifies the new way of thinking adopted by the lady in question before she died.

The split constituent offers a definite description of a referent that has not been talked about before. Thus, it is not anchored to a prior mention, but rather, is identifiable through the cultural knowledge the speaker assumes that she and the co-participant share (on the expressions of identifiability in Finnish, see Laury 1997). The split and the hesitation can be seen as indications that the speaker may see potential problems with the identifiability of the referent.

Thus, it is quite rare that NPs split into two intonation units. But if an NP was produced in two or more intonation units, it was more likely to function in the object role (about half of the splits were in the object role) than any other role (cf. ex. 52 and 53). In order to interpret this finding it is important to bear in mind the fact that only constituents that consist of several parts can split. As we have seen in Chapter 2, pronominal NPs or zeros are preferred over full NPs in the subject role, whereas in the object and the oblique role the proportion of full NPs is higher. It is not entirely clear to me why NPs in the object role should lend themselves to splits more readily than obliques. One possible explanation is that obliques primarily serve to orient to the settings whereas object NPs more often serve to track referents. As we have seen, the split NPs often occur in contexts where there is some problem either in the choice of the referential term (cf. ex. 53) or in characterizing the referent (ex. 52).

5.1.3 Summary

In conclusion, intonation unit boundaries almost always occur at constituent boundaries. NPs are rarely split across intonation units, and even in cases where a constituent is split across more than one intonation unit the splits are not “errors” due to performance factors but show clear patterns: the splits occur between the determiner and the head, rather than e.g. between a characterizing modifier and the head. Most of the time intonational phrasing is convergent with syntactic phrase structure. However, as we will see in the next section, intonational patterning does not lend unequivocal support to the category of verb phrase.

5.2 Intonational grouping and clausal organization

As was discussed in the beginning of this chapter, the view that intonation units correspond to clauses (or sentences, depending on definitions) seems to be quite widely accepted. However, it is usually assumed rather than argued for. As we saw in Chapter 2 (esp. Section 2.1), it is often far from clear what is meant by the notion of clause. In this section, I would like to discuss how the approach to clauses and clause structure that was developed in Chapter 2 could be further strengthened by studying the interplay between intonational grouping and structuring of clausal syntactic relations.

Many transcription theories employ the notion of clause or sentence in describing the transcription conventions of intonation contours. Thus, they speak about sentence-final intonation that is marked with a period and question intonation marked with a question mark (see e.g. Chafe 1994). Couper-Kuhlen and Selting (1996:41) see a potential danger here: “[T]he use of punctuation marks, for instance, for the representation of phrase-final intonation depends on stereotypical associations between sentence-type and intonation which are misleading.”

In Chapter 2, I proposed the following definition for clauses: clauses are emergent syntactic units that consist of the predicate and its core arguments which together make up the clause core. The core may be accompanied by additional adjuncts, but the adjuncts do not constitute the clause in the same way as the predicate and its core arguments do. In other words, clausal syntactic relations are clearest in the clause core, whereas peripheral arguments often give information about circumstantial relations which are syntactically not tightly

related to the clause core. It would seem reasonable to hypothesize that they are intonationally not necessarily bound to the core, but may just as well occur in a different intonation unit from the core.

In what follows, I will first discuss the integration of oblique NPs into clausal structure. Then there will be a discussion of the distribution of clause cores across intonation units. Finally, I will briefly discuss how intonation can be used to make projections of the continuation of the interaction.

5.2.1 Oblique NPs and clause structure

In this section, I will show that it is not always possible to tell which clause a given oblique NP belongs to, and that it may be that this question is not crucial at all. Instead, I will propose that oblique NPs show a continuum of integration, with clear clausal arguments at one end, and free NPs at the other end.

Example (54) is part of a series of stories about cats. It is preceded by a story about a cat who used to jump down from the balcony of a 3rd floor apartment. Earlier there had been a story about a little girl who had dropped a cat from the balcony of a 5th floor apartment. In the preceding stories, the cats have faced all kinds of dangers but always remained unscathed. At this point, Raita offers a different story:

(54) Birthday

1 Raita: *mut sit se Onni-n kissa teki se-n itsari-n.*
 but then 3SG O.-GEN cat committed 3SG-ACC suicide-ACC
 But Onni's cat (that cat of Onni) committed suicide.

2 Leena: ... *nii.*
 Right.

3 Noora: *hä?*
 Huh?

4 Raita: *Kruunuhaa-ssa.* ←
 Kruununhaka-INE
 In Kruununhaka.

5 Leena: [₁ *Onnin ja* ₁] *Pirkon kissa.*
 Onni and Pirkko's cat.

6 Raita: [₁ *mä olin-* ₁]
 I was—

7 Raita: ... *olik se Ossi vai mikä sen nimi oli.*
 Was it Ossi or what the name was.

- 8 Noora: *voi kun ihana nimiki* Os_[2SGi.2]
Oh even a cute name Ossi.
- 9 Sanna: _[2mhm.2]
- 10 Mella: _[2mm.2] jo_[3O=.3]
Mhm. Yeah.
- 11 Raita: _[3se oli 3]
It was—
- 12 Leena: _[3mutta=3] se oli
but it was
- 13 *eik se ollu Onni ja Pirkko ajanu sen s- kissan*
wasn't it Onni and Pirkko who had driven the th- cat
- 14 *siihen tilantee* _[4seen.4]
into that situation.
- 15 Raita: _[4niin.4]
Right.
- 16 Raita: *neljäs tai viides kerros,*
fourth or fifth floor,
- 17 *kumpi se oli.*
which one was it.
- 18 Leena: (TSK) ^Viron-kadu-_[5lla.5] ←
Viro-street-ADE
On Viro street.
- 19 (X): _[5^nii.5]
Right.
- 20 Raita: *joo Viron-kadu-lle sitte* _[6niin.6]ku, ←
yeah Viro-street-ALL then like
Yeah. (On)to Viro street then like...
- 21 Sanna: _[6mm.6]
Mhm.
- 22 *voi men-nä miele-n-terveys*
can go-INF mind-GEN-health
You may lose your mental health
- 23 *jos muutta-a* <INH Viron-kadu-lle INH>?
if move-3SG Viro-street-ALL
if you move to Viro street.

I will first analyze the example in order to show how it is built up syntactically and interactionally, and then I will focus on the oblique NPs in the sequence in terms of whether they relate to clausal constructions or form units of their own.

The example starts out with Raita's dramatic opening line, which summarizes the main point of the story: a cat committed suicide (line 1).² In line 2, Leena responds to this with the particle *nii* 'Right', produced with a falling terminal contour (marked with a period in the transcript). According to Sorjonen (1997:426), the particle *niin* is used in story-telling contexts at certain key points, inter alia when the teller has introduced the main characters. She also shows that *niin*, when produced with a falling terminal contour as an intonation unit of its own (like in ex. 54), claims recognition of a referent referred to by the co-participant (Sorjonen 1997:427). Thus, with her response in line 2, Leena displays that she knows what Raita is going to tell about. Furthermore, Sorjonen (1997:313–314) shows that the particle *niin* can be used in response to a mutually known issue to display affiliation. Noora aligns herself as a recipient of the story with her response *häh* 'huh' requesting more information. In line 4, Raita adds the inessive NP *Kruunuhaassa* 'in Kruununhaka' to pin down the event described in line 1 to a specific place. In line 5, Leena repairs Raita's line 1 by repeating and expanding the NP *se Onnin kissa* by adding another genitive modifier to it *Onnin ja Pirkon kissa* 'Onni and Pirkko's cat'. Raita then adds more information about the cat by naming it (line 7). In lines 12–14, Leena gives a psychologizing evaluation of the situation leading to the event, and Raita responds with an affiliative *niin* (line 15; cf. line 2). According to Sorjonen (1997:314), the particle *niin*, when produced as an intonation unit of its own, is usually followed by a continuation by the same speaker. Here, Raita continues to add further details about the event (lines 16 and 17; cf. line 2 which is not followed by a continuation by the same speaker). This is done with a free NP *viides tai kuudes kerros* 'fifth or sixth floor' (line 16). The free NP in line 16 *viides tai kuudes kerros* 'fifth or sixth floor' makes a predication about the apartment by characterizing it (see Section 4.3.1), but note that the apartment

2. The opening line is interesting syntactically: it contains a construction with two full NP arguments (*se Onnin kissa* 'that cat of Onni' and *sen itsarin* 'the suicide'). As we saw in Chapters 2 and 3, this is very rare in spoken discourse in general as well as in the present data. In the Birthday-conversation, there is a series of stories about cats and example (54) is one them. Interestingly enough, all the stories in this series start with a construction with two full NP arguments. It could be argued that this syntactically marked construction type is used as a structural device to mark the stories as belonging to the same series (see Helasvuo 1991).

itself is not mentioned at all in the whole story. In line 18, Leena specifies the exact place where it all happened with an adessive NP *Vironkadulla* ‘on Viro street’. Raita responds to this with an acknowledging *joo* and continues by repeating the same NP in a different case form, the allative *Vironkadulle* ‘onto Viro street’ (line 20). We will discuss these two NPs in more detail shortly. The closing lines 22–23 by Sanna pursue further the personifying and psychologizing line of approach created by Raita (line 1) and continued by Leena (lines 12–14), but at the same time, they look at the contents of the story from a more general perspective.

In example (54) there are three oblique NPs that are in intonation units where there is no predicate verb (lines 4, 18 and 20). I will now discuss the analysis of these obliques. In the example, there is a transitive predication in line 1, and the oblique NP in line 4 locates the event described in line 1 in a specific place. The oblique NP is added as an increment to the clause in line 1. Note that the clause in line 1 is syntactically complete in that it does not need a continuation, and also, there is a final falling intonation contour (marked with a period) in the end of the clause. Nevertheless, it is possible to add an increment to it, thus extending the construction (Ford, Fox and Thompson 1997). We could analyze the oblique NP (line 4) as an adjunct of the clause in line 1. It is important to note that the intervening talk — the two response tokens in lines 2 and 3 — does not start anything new, but is oriented to the same activity.

In lines 18 and 20, there are two locative NPs which are different case forms (adessive and allative) of the same word, *Vironkatu* ‘Viro street’. The two case forms focus on different aspects of the event: the adessive NP (line 18) locates the whole event (the suicide Onni’s cat has committed), whereas the allative NP (line 20) makes a predication about the cat and its movement onto the street (see Section 4.3.2). However, there is no verbal predication in the context about the cat and how it got onto the street, whether it jumped or fell. The event has been framed as a suicide (cf. line 1), which gives an agentive interpretation to the event.

Although the two obliques in lines 18 and 20 are similar in terms of morphological form, they differ in terms of their syntactic independence and interactional contribution. Let us first look at the adessive NP *Vironkadulla* ‘on Viro street’ in line 18. There are (at least) three possible ways to analyze it:

- i. It is syntactically compatible with the clause in line 1, and thus, could be analyzed as being part of that clause. However, there is intervening talk that focuses on different things, such as the background situation leading to the suicide (lines 12–14) and details of the apartment (lines 16–17). Therefore, it seems somewhat far-fetched to analyze it as being part of the clause in line 1.

- ii. One could also argue that with the adessive NP, Leena continues Raita's talk by adding to the details of the apartment. In this interpretation, both the free NP and the adessive NP would be syntactically unattached.
- iii. Thirdly, we could analyze it as an increment to the preceding discussion, and as such, not necessarily tied to any particular construction, and thus, syntactically free. This is possible because there is a strong meaning potential in the adessive NP itself, brought by the semantics of the case ending and the lexical semantic content of the NP.

In line 20, the oblique NP is in the allative case, and it is syntactically not compatible with any of the clausal predications in the preceding sequence. It makes a predication about the cat: the allative case indicates movement onto something, namely the referent of the NP to which the allative ending is attached (Viro street), and the understood theme of the predication (the one which the predication is about) is the cat. This third oblique NP differs from the previous ones (lines 4 and 18) in one important respect: unlike the two preceding ones which are adding to previous predications, it contributes with a predication of its own.

The point of my rather lengthy discussion of example (54) is that in the analysis of obliques, we do not always have clear syntactic or semantic evidence that a given oblique phrase belongs to one clause rather than another in the context. Of the oblique NPs in example (54), the first one (line 4) is most clearly attached to a clause, whereas for the second one (line 18), this is less clearly the case. The last one (line 20) is clearly unattached.

In conclusion, I have shown that often enough, the data does not give basis for a strong argument for the analysis of a given oblique NP as part of one clause or the other. Thus, oblique NPs show a certain vagueness in regard to their syntactic integration. It is no wonder that they often form intonation units of their own. Because of this vagueness, I have decided to exclude obliques when I studied the distribution of clauses across intonation units, but instead, I only looked at clause cores.

5.2.2 Intonational grouping and the clause core

The data show a clear tendency for clause cores to be produced in a single intonation unit. In this section, we will study this tendency more closely. We will also look at deviations of this general pattern, and discuss the implications of the findings.

If we study the clause in terms of the predicate and its core arguments, it is most often the case that the predicate and its core arguments are in the same intonation unit. This was the case in 90.5% of the clauses in my data. In the English data of Helasvuo (1997a), the predicate and its core arguments appeared in the same intonation unit in 95% of the clauses, thus showing a similar tendency as the Finnish data.

The data show that if either or both of the core arguments were in a different intonation unit from the predicate, it was usually the object that was in a different intonation unit from the predicate and the subject. In the Finnish data, there were 46 objects (out of a total of 297 objects) but only 13 subjects (out of a total of 726 subjects) that were in a different intonation unit from the verb. This contradicts Cruttenden's claim (see the beginning of this chapter) that intonation unit boundaries should occur between the subject and the predicate, rather than between the predicate verb and the object.

As was discussed in the beginning of this chapter, Langacker (1997) has proposed a model of constituency that allows for different kinds of groupings. Thus, in a clause such as *Alice likes Bill* it is possible that the component structures *likes* and *Bill* form the composite structure *likes Bill*, which is then integrated at a higher level with *Alice* to form *Alice likes Bill*, or alternatively, *Alice* and *likes* form a composite *Alice likes*, with which *Bill* is then ingrated to form *Alice likes Bill*. Langacker notes, however, that the alternative groupings are not equally natural; he says, "I presume that the classic S (VO) constituency is unmarked, representing the default in English, but that alternate patterns involving an SV constituent are used when there is structural motivation to resort to them." On the basis of the English data of Helasvuo (1997a) and the present data, intonational grouping does not provide support for the supposed primacy of the VO-grouping.

Furthermore, Helasvuo (in press) and Scheibman (forthcoming) show that 1st person subjects tend to co-occur with mental verbs (see also Kang 1997 for French) and that they together form regular and frequent combinations which may even crystallize into one fixed unit, such as *I think, you know, I mean* etc. In English, these subject + verb-combinations have been grammaticized as discourse particles (see Kärkkäinen 1998, Thompson and Mulac 1991, Östman 1981). In an in-depth study of the grammaticization of *I think*, Kärkkäinen (1998) shows that *I think* has been grammaticized into a marker of epistemic stance and that it forms a unit that is produced either before or after the proposition it is modifying. In sum, these grammaticization tendencies can be seen as providing further support for an SV-grouping on a par with a VO-grouping.

Interestingly enough, those clauses in which the clause core is distributed across several intonation units show features of low transitivity (in the sense of Hopper and Thompson 1980; see Section 2.4.3). They tend to show imperfective aspect, describe non-punctual actions or processes with no permanent effect on the object. Often they involve cognitive or communicative verbs. The object is usually not highly individuated. Consider example (55).

(55) Birthday

- 1 Noora: @ *kissa istu siel* ^*lihatiski-n* ääressä,
cat sat there meat counter-GEN at
The cat was sitting there at the meat counter (in a super-
market)
- 2 (H) *katto* [^*sil*]*mä-t ymmyrkäis-i-nä*,
looked eye-PL rounded-PL-ESS
looking, with its eyes wide open,
- 3 Sanna: [(H)]
- 4 Noora: ^*kano-j-a* grilli-ssä.
chicken-PL-PTV barbecue-INE
at chickens in the barbecue.

In example (55), the predicate verb *katto* ‘was looking at’ (in line 2) is a cognitive verb describing a process. It is followed by an adverbial construction *silmät ymmyrkäisinä* ‘with its eyes rounded’ describing the subject (*kissa* ‘cat’, line 1).³ Only after that does the speaker produce the object *kanoja* ‘chickens-PTV’ in a different intonation unit (line 4). The object NP stands in the partitive case, which was shown to be strongly associated with low transitivity (Section 2.4.3). As is typical of partitive NPs, the referent is not highly individuated.

Hopper and Thompson (1980) suggest that transitivity is related to grounding in discourse so that clauses exhibiting high transitivity are foregrounded in discourse whereas clauses of low transitivity are backgrounded. Although there may be problems in applying the notion of grounding as a manifestation of transitivity to conversational discourse (for discussion see e.g. Helasvuo 1991),

3. It is also noteworthy that the adverbial construction *silmät ymmyrkäisenä* ‘with its eyes rounded’, together with the predicate verb *katto* ‘look at, watch’, expresses a certain way of watching or looking, which is very intense. This could be taken to suggest that they form a verbal expression together, here expressed as one intonation unit, and the object *kanoja* ‘chickens’, which brings new information, occurs in an intonation unit of its own (Wallace Chafe, pc).

prosody is certainly an important factor in the structuring of larger stretches of discourse, especially narrative discourse (see Chafe 1980, Kumpf 1987).

The intonation patterns found in the data offer us an interesting perspective on grammatical structuring: in terms of intonational groupings, the bond between the subject and the verb seems to be even stronger than that between the verb and the object, because objects are more often in a different intonation unit from the verb than are subjects. In Section 2.6, I showed that the mutual ordering of subject and verb is very stable with the subject usually preceding the verb. This tendency is very robust, especially in clauses with personal pronouns functioning as subjects. Thus, subjects and verbs show fixed word order patterns. Therefore, it is not surprising to find that subjects are more often presented in the same intonation unit with the predicate than are objects.

From the point of view of Finnish linguistics, it is not surprising to find that objects are more often in a different intonation unit from the predicate than are subjects. In fact, many researchers have cast doubt on the reality of a verb phrase in Finnish (most recently Vilkuna 1996). For example, the verb and the object do not necessarily form a continuous constituent. Example (55) illustrates this: between the predicate verb *katto* ‘was looking’ (line 2) and the object *kanoja* ‘chicken’ (line 4), there is an adverbial construction *silmät ymmyrkäisinä* ‘with its eyes rounded’ (line 2) that characterizes the referent of the subject NP *kissa* ‘cat’ (line 1).

There is evidence from other languages that the category of verb phrase may be problematic. Fox and Jasperson (1995: 123) note in their study on the syntax of repair in English that the category verb phrase “may not be oriented to in repair in English conversation”: if the repair is initiated postverbally, the repairing segment does not recycle back to the verb. Fox and Jasperson suggest (1995: 127) that this “perhaps indicates that the notion of verb phrase is syntactically problematic” for English (see also Hopper 1997). It seems that the verb phrase does not emerge as a strong, “classic” constituent in the sense of Langacker (1995, 1997).

In this section, I have shown that there is a strong tendency in spoken discourse for clause cores to be presented in a single intonation unit. However, if an argument was produced in a different intonation unit from the predicate, it was much more likely to be the object than the subject. This indicates that in terms on intonational grouping, the bond between the subject and the verb seems even stronger than that between the verb and the object.

5.3 Summary

In this chapter, I have shown that intonation is an important means of combining elements together to form units, on a par with grammatical and semantic grouping. Most often the different groupings yield similar results. We found that in the Finnish data, intonation unit boundaries almost always occur at constituent boundaries. I further showed that even in cases where a constituent is split into subsequent intonation units the splits are not random “errors” due to performance factors but show clear patterns, which illustrate how reference is created and negotiated in discourse.

In the data, there was a strong tendency for clause cores to be presented in a single intonation unit. I see this as a strong argument for clauses as emergent constructions consisting of the predicate and its arguments. They may be accompanied by optional adjuncts, which may or may not appear in the same intonation unit as the core. In this chapter we have seen that the clause core shows strong intonational unity. I take this to provide evidence for my claim in Chapter 2 that syntactic relations appear most clearly in the clause core. The data analysis shows the robustness of the category of NP: as we saw in the analysis, NPs that are clausal constituents are rarely split into two intonation units. Free NPs form syntactic units of their own, and they are most often also produced in intonation units of their own.

CHAPTER 6

Conclusion

We started out with the question of what clauses are, how they are realized in discourse, and what other kinds of recurrent syntactic units there are besides clauses. I have sought to answer these questions by looking at discourse data from conversational Finnish. The answers are of course language-specific, and sometimes even genre-specific, but at the same time, they have cross-linguistic implications which I have hinted at and which we will discuss shortly.

In light of the Finnish discourse data, it is the clause core where grammatical relations emerge most clearly. The clause core emerges as a unit not only in terms of syntactic coding, but also intonationally. Furthermore, the clause core was shown to play a central role in managing information, *inter alia* in delivering new information or tracking participants in discourse. We have also seen strong evidence for the robustness of the category of noun phrase, both as a clausal constituent and as a free syntactic unit, the free NP.

We have discussed the coding of grammatical relations in terms of case marking, agreement and word order. All three coding systems highlight the role of the clause core as the locus for the most explicit coding. First, agreement serves to code the relationship between the subject and the predicate verb, and in negative terms, also the object. Secondly, the most robust tendencies in word order patterns can be seen in the relationship between the subject and the verb. And finally, case marking shows interesting interactions in the core roles: for example, if there is no nominative subject in the clause, the object can stand in the nominative. In the oblique cases, however, there is no such interaction. Thus, all three coding strategies discussed here point to the clause core as the locus of the emergence of syntactic relations.

In Chapter 2, I argued that the category of person provides the key to understanding how grammatical coding systems work in Finnish. As I have shown, the marking of personal pronouns provides a new perspective on case marking of syntactic relations in the clause core, especially the marking of the object. To describe object marking, I took personal pronouns as a starting point. As I have shown, personal pronouns exhibit a straight nominative-accusative coding system, where the subject is unmarked for case, but the

object is obligatorily marked (either with the accusative or the partitive). The marking of full NP objects is more complicated, but the alternations can be explained more easily on the basis of how the personal pronouns are marked. In the agreement system, the agreement patterns are most explicit in the 1st and 2nd person, but in the 3rd person there is more indeterminacy. And finally, word order in clauses where the subject is a personal pronoun shows a fixed Subject–Verb ordering, whereas in clauses with a full NP subject, there is more flexibility.

Thus, personal pronouns show the most explicit coding of grammatical relations in the clause core. This is a finding that is specific to Finnish; however, it does have cross-linguistic implications, which can only be investigated by carefully studying conversational discourse data from different languages.

In previous research, it has been acknowledged that personal pronouns show coding patterns that are different from those of other pronouns and full NPs. Most notably, Silverstein (1976, 1981) has shown that if a language shows split patterns in coding, personal pronouns are most likely to exhibit nominative-accusative marking. The fact that personal pronouns are coded differently has often led to the exclusion of pronouns from the description of coding systems. In other words, the coding of personal pronouns has been marginalized. Even studies that use actual discourse data are sometimes biased in this respect if they are based on third person narrative only. I have shown that in conversational discourse, 1st and 2nd person pronouns figure significantly in the subject role, especially in transitive clauses. I argue therefore that personal pronouns should be taken as a central force in the structuring of argument relations in discourse. The particular ways in which personal pronouns are used in Finnish discourse are of course not readily generalizable to other languages, but I do assume that personal reference is important in conversational language in general. We are clearly in need of empirical cross-linguistic studies on the role of the category of person in the coding of grammatical relations in conversational discourse.

I have shown that in Finnish personal pronouns realize a straight nominative-accusative system with a clearly distinguished subject role. In terms of the grammaticization of the subject role, we could say that personal pronouns exhibit a highly grammaticized nominative-accusative coding system, and, accordingly, a clearly defined subject role. In contrast, the marking of full NPs is closer to a neutral alignment system, where the coding does not clearly distinguish the subject as a distinct grammatical role, but instead, there is

much overlap in the coding system. I argue nevertheless that full NPs can be recognized as subjects based on the analogical model of pronominal subjects.

In addition to their role as clausal constituents, noun phrases can function as free syntactic units, the so-called free NPs. I have shown that free NPs can be used in a variety of functions ranging from typical noun functions relating to the creation of reference to less prototypical ones such as orientation and classification. I have shown that speakers can use free NPs as a resource in negotiating reference, but also to make predications. They can be used to help identify an entity or to focus on referent that is a member of a larger set of referents already under discussion. Predicating free NPs can serve to characterize referents or to disambiguate reference; i.e. the participants can use them to check their understanding of prior talk. Using constructions with a theme and orientation speakers can make predications that initiate something new, rather than looking back to prior discourse.

As further evidence for the clause as an emergent syntactic unit, we found in Chapter 5 that the clause core shows strong intonational unity. Furthermore, the intonation analysis showed the robustness of the category of NP as NPs are rarely split into two intonation units. Although there was a strong tendency for the clause core to be produced as one intonation unit, there were nevertheless cases where one of the core arguments was in a different intonation unit from the predicate. Interestingly enough, if an argument was produced in a different intonation unit from the predicate, it was more likely to be the object than the subject. This patterning suggests a subject–verb grouping on a par with the traditional verb–object grouping. Thus, there is evidence in the data for a more flexible analysis of grammatical constituency than has traditionally been assumed, allowing for different kinds of groupings among the elements.

Thus, there is strong evidence for the clause as an emergent grammatical unit in Finnish. By no means would I like to suggest that this would be true of languages in general. In fact, we have discourse studies from other languages that point to the fact that this is not the case in all languages. Inter alia, one of the basic tenets of Tao (1996) is that the clause does not emerge so clearly as a level of syntactic organization in Mandarin discourse, but instead, syntactic relations emerge clearest on the phrase level. Furthermore, Nakayama (1997) shows that in Nuu-chah-nulth (Nootka), syntactic structure is, for the most part, not anchored to grammaticized categories and markers, but is realized in the systematic arrangements of words. In other words, syntax emerges on the word level rather than in agreement patterns or case marking as in Finnish.

From a grammaticization perspective we could say that languages resort to very different strategies in trying to handle the various discourse needs that speakers are faced with, such as the need to express events or track participants of an event. We need careful discourse studies of individual languages in order to build a cross-linguistic understanding on how grammar works.

References

- Andrews, Avery 1985: The Major Functions of the Noun Phrase. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Vol. 1: 62–154. Cambridge University Press, Cambridge.
- Ashby, William J. 1988: The Syntax, Pragmatics, and Sociolinguistics of Left- and Right-Dislocations in French. *Lingua* 75: 203–229.
- Ashby, William J. 1992: The variable use of *on* versus *tu/vous* for indefinite reference in Spoken French. *French Language Studies* 2: 135–157.
- Ashby, William J. 1995: French Presentational Structures. In John Amastae et al. (ed.), *Selected Proceedings of XXII Linguistic Symposium on Romance Languages*. John Benjamins, Amsterdam.
- Ashby, William & Paola Bentivoglio 1993: Preferred Argument Structure in Spoken French and Spanish. *Language variation and change* 5: 61–76.
- Auer, Peter 1996: On the prosody and syntax of turn-continuations. In Elizabeth Couper-Kuhlen & Margret Selting (eds), *Prosody in conversation*. Pp. 57–100. Cambridge University Press, Cambridge.
- Bentivoglio, Paola 1992: Linguistic Correlations between Subjects of One-Argument Verbs and Subjects of More-Than-One-Argument Verbs in Spoken Spanish. In Paul Hirschbühler & Konrad Koerner (eds), *Romance Linguistics and Modern Linguistic Theory*. Current Issues in Linguistic Theory, vol. 91. Pp. 11–24. John Benjamins, Amsterdam.
- Blake, Barry & Graham Mallinson 1981: *Language Typology: Crosslinguistic Studies in Syntax*. North-Holland, Amsterdam.
- Blake, Barry 1994: *Case*. Cambridge University Press, Cambridge.
- Brown, Gillian 1977: *Listening to Spoken English*. Longmans, London.
- Bruun, Kettil 1958: *Drinking behaviour in small groups*. The Finnish Foundation for alcohol studies 9, Helsinki.
- Bybee, Joan, Revere Perkins & William Pagliuca 1994: *The Evolution of Grammar: Tense, Aspect, and Modality in the languages of the world*. The University of Chicago Press, Chicago.
- Bybee, Joan & Joanne Scheibmann 1997: The effect of usage on degrees of constituency: the reduction of don't in English. Paper presented at the Conference on Constituency and Discourse, May 1997. University of California, Santa Barbara.
- Chafe, Wallace L. 1976: Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In Charles N. Li (ed.), *Subject and Topic*. Pp. 25–55. Academic Press, New York.
- Chafe, Wallace L. 1979: The flow of thought and the flow of language. In Talmy Givón (ed.), *Discourse and syntax*. Syntax and semantics, vol. 12. Pp. 159–181. Academic Press, New York.

- Chafe, Wallace L. 1980: The Deployment of Consciousness in the Production of a Narrative. In W. Chafe (ed.), *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Pp. 9–50. Ablex, Norwood, NJ.
- Chafe, Wallace L. 1987: Cognitive Constraints on Information Flow. In Russell Tomlin (ed.), *Cohesion and Grounding in Discourse*. Pp. 21–51. John Benjamins, Amsterdam.
- Chafe, Wallace L. 1994: *Discourse, Consciousness, and Time. The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press, Chicago.
- Chafe, Wallace L. & Marianne Mithun 1995: What use are S, A, and O? Paper given in the conference on Functional Approaches to Grammar, July 1997. University of New Mexico, Albuquerque, NM.
- Comrie, Bernard 1981: *Language Universals and Linguistic Typology*. Blackwell, London.
- Couper-Kuhlen, Elizabeth 1996: The prosody of repetition: on quoting and mimicry. In Elizabeth Couper-Kuhlen & Margret Selting 1996a. Pp. 366–405.
- Couper-Kuhlen, Elizabeth & Margret Selting (eds) 1996a: *Prosody in Conversation*. Cambridge University Press, Cambridge.
- Couper-Kuhlen, Elizabeth & Margret Selting 1996b: Towards and interactional perspective on prosody and a prosodic perspective on interaction. In Elizabeth Couper-Kuhlen & Margret Selting 1996a. Pp. 11–56.
- Couper-Kuhlen, Elizabeth & Margret Selting (eds.) (in press): *Interactional Linguistics*. John Benjamins, Amsterdam.
- Croft, William 1991: *Typology and Universals*. Cambridge University Press, Cambridge.
- Croft, William 1995: Intonation units and grammatical structure. *Linguistics* 33: 839–882.
- Croft, William 1997: Intonation units and grammatical structure in Wardaman and English. Paper given in the symposium on Constituency and Discourse, May 1997. University of California, Santa Barbara.
- Cruttenden, Alan 1986: *Intonation*. Cambridge University Press, Cambridge.
- Crystal, David 1969: *Prosodic Systems and Intonation in English*. Cambridge University Press, Cambridge.
- Cumming, Susanna 1984: The sentence in Chinese. *Studies in Language* 8.3: 365–395.
- Dahl, Östen 1979: Typology of Sentence Negation. *Linguistics* 17: 79–106.
- Davis, Philip (ed.) 1996: *Alternative Linguistics. Descriptive and Theoretical Modes*. John Benjamins, Amsterdam.
- Dixon, R. M. W. 1972: *The Dyirbal Language of North Queensland*. Cambridge University Press, Cambridge.
- Dixon, R. M. W. 1977: *A Grammar of Yidiny*. Cambridge University Press, Cambridge.
- Dixon, R. M. W. 1979: Ergativity. *Language* 55: 59–138.
- Du Bois, John W. 1980: Beyond Definiteness: The Trace of Identity in Discourse. In Wallace Chafe (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Ablex, Norwood, NJ.
- Du Bois, John W. 1985: Competing Motivations. In John Haiman (ed.), *Iconicity in Syntax*. John Benjamins, Amsterdam.
- Du Bois, John W. 1987a: The Discourse Basis of Ergativity. *Language* 63: 805–855.
- Du Bois, John W. 1987b: Absolutive zero: Paradigm Adaptivity in Sacapultec Maya. *Lingua* 71: 203–222.

- Du Bois, John W., Stephan Schuetze-Coburn, Danae Paolino & Susanna Cumming 1991: Discourse Transcription. Ms. Department of Linguistics, University of California, Santa Barbara.
- Du Bois, John W., Stephan Schuetze-Coburn, Danae Paolino & Susanna Cumming 1993: Outline of Discourse Transcription. In Jane A. Edwards & Martin D. Lampert (eds.), *Talking Data: Transcription and Coding in Discourse Research*. Lawrence Erlbaum, Hillsdale, N.J.
- Du Bois, John W. & Sandra A. Thompson 1990: Dimensions of a Theory of Information Flow. Ms. University of California, Santa Barbara.
- Du Bois, John W., William Ashby & Lorraine Kumpf (eds.) forthcoming: *Preferred Argument Structure: Grammar as Architecture for Function*. John Benjamins, Amsterdam.
- Duranti, Alessandro 1994: *From Grammar to Politics. Linguistic Anthropology in a Western Samoan Village*. University of California Press, Los Angeles.
- Durie, Mark 1988: Preferred Argument Structure in an Active Language. *Lingua* 74: 1–25.
- Durie, Mark 1994: A Case Study of Pragmatic Linking. *Text* 14 (4): 495–529.
- Durie, Mark (forthcoming): New light on information pressure. Information conduits, “escape valves”, and role alignment stretching. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred Argument Structure: Grammar as Architecture for Function*. John Benjamins, Amsterdam.
- Erringer, Anu 1996: The functions of demonstrative adjectives *semmoinen*, *tämmöinen* and *tuommoinen* in Finnish conversations. MA Thesis, Department of linguistics, University of Colorado, Boulder.
- Foley, William & Robert Van Valin Jr. 1984: *Functional Syntax and Universal Grammar*. Cambridge University Press, Cambridge.
- Ford, Cecilia 1993: *Grammar in Interaction: Adverbial Clauses in American English Conversations*. Cambridge University Press, Cambridge.
- Ford, Cecilia, Barbara Fox & Sandra A. Thompson 1997: Increments in Conversation and their Relevance to Constituency. Paper given at the Workshop on Constituency and Discourse. University of California, Santa Barbara.
- Ford, Cecilia & Sandra A. Thompson 1996: Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns. In Elinor Ochs, Emanuel Schegloff & Sandra A. Thompson (eds.), *Interaction and Grammar*. Pp. 134–184. Cambridge University Press, Cambridge.
- Forsberg, Hannele 1994: Havaintoja kysymystyyppistä *otat sie ~ sä*. *Virittäjä* 98: 60–68.
- Fox, Barbara 1987: *Anaphora and the Structure of Discourse*. Cambridge University Press, Cambridge.
- Fox, Barbara, Makoto Hayashi & Robert Jaspersion 1996: Resources and repair: a cross-linguistic study of syntax and repair. In Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and grammar*. Studies in Interactional Sociolinguistics 13. Pp. 185–237. Cambridge University Press, Cambridge.
- Fox, Barbara & Robert Jaspersion 1995: A syntactic exploration of repair in English conversation. In Philip W. Davis (ed.), *Alternative linguistics: Descriptive and theoretical modes*. CILT 102. Pp. 77–134. Benjamins, Amsterdam.

- Franck, Dorothea 1985: Sentences in conversational turns: A case of syntactic 'double bind'. In Marcelo Dascal (ed.), *Dialogue: An interdisciplinary approach*. Pp. 233–245. John Benjamins, Amsterdam.
- Geluykens, Ronald 1987: Tails (right-dislocations) as a Repair Mechanism in English Conversation. In J. Nuyts & G. de Schutter, *On Word Order and Functional Grammar*. Pp. 119–129. Foris, Dordrecht.
- Geluykens, Ronald 1988: The Interactional Nature of Referent-Introduction. *Papers from the 24th Regional Meeting of the Chicago Linguistic Society*. Pp. 141–154. Chicago Linguistic Society, Chicago.
- Geluykens, Ronald 1992: *From Discourse Process to Grammatical Construction: On Left-Dislocation in English*. John Benjamins, Amsterdam.
- Givón, Talmy 1979: From Discourse to Syntax: Grammar as a Processing Strategy. In Talmy Givón (ed.), *Discourse and Syntax*. Syntax and Semantics, vol. 12. Academic Press, New York.
- Givón, Talmy (ed.) 1983: *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. John Benjamins, Amsterdam.
- Givón, Talmy 1984: *Syntax*. Vol. 1. John Benjamins, Amsterdam.
- Givón, Talmy 1991: Serial verbs and the mental reality of 'event': grammatical vs. cognitive packaging. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*. Vol. 1. Pp. 81–128. John Benjamins, Amsterdam.
- Givón, Talmy 1994: *A function-based introduction to English grammar*. John Benjamins, Amsterdam.
- Goffman, Erwin 1981: *Forms of talk*. Basil Blackwell, Oxford.
- Goodwin, Charles 1979: The Interactive Construction of a Sentence in Natural Conversation. In G. Psathas (ed.), *Everyday Language: Studies in Ethnomethodology*. Pp. 97–121. Erlbaum, New York.
- Goodwin, Charles 1981: *Conversational Organization. Interaction between Speakers and Hearers*. Academic Press, New York.
- Goodwin, Charles & Marjorie Harness Goodwin 1987: Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments. *IPRA Papers in Pragmatics* 1:1–54.
- Haiman, John 1983: Iconic and Economic Motivation. *Language* 59: 781–819.
- Haiman, John (ed.) 1985a: *Iconicity in Syntax*. John Benjamins, Amsterdam.
- Haiman, John 1985b: *Natural Syntax: Iconicity and Erosion*. Cambridge University Press, Cambridge.
- Halliday, M. A. K. 1967: Intonation and Grammar in British English. *Janua Linguarum, Series Practica* 48. The Hague, Mouton.
- Halliday, M. A. K. 1985: *Introduction to Functional Grammar*. Edward Arnold, London.
- Hakanen, Aimo 1972: Normaalilause ja eksistentiaalilause. (Referat: Normalsatz und Existentialsatz.) *Sananjalka* 14: 36–76.
- Hakulinen, Auli 1983: Subjektin kategoria vai nominaalijäsenten subjektimaisuus. In Auli Hakulinen & Pentti Leino (eds.), *Nykysuomen rakenne ja kehitys* 1. Pp. 238–251. Finnish Literature Society, Helsinki.
- Hakulinen, Auli 1989: Palauteilmauksista. In Auli Hakulinen (ed.), *Suomalaisen keskustelun keinoja* 1. Kieli 4: 98–114. Department of Finnish, University of Helsinki, Helsinki.

- Hakulinen, Auli 1993: The grammar of opening routines. In Susanna Shore & Maria Vilkkuna (eds.), *SKY 1993 Yearbook of the Linguistic Association of Finland*. Pp. 149–170. Helsinki.
- Hakulinen, Auli & Fred Karlsson 1979: *Nykysuomen lauseoppia*. Finnish Literature Society, Helsinki.
- Hakulinen, Auli, Fred Karlsson & Maria Vilkkuna 1980: *Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus*. Publications of the Department of General Linguistics, vol. 6. University of Helsinki, Helsinki.
- Hakulinen, Lauri 1979: *Suomen kielen rakenne ja kehitys*. 4th edition. Otava, Keuruu.
- Harris, Alice C. & Lyle Campbell 1995: *Historical Syntax in Cross-Linguistic Perspective*. Cambridge Studies in Linguistics 74. Cambridge University Press, Cambridge.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer 1991: *Grammaticalization. A Conceptual Framework*. University of Chicago Press, Chicago.
- Heinämäki, Orvokki 1983: Aspect in Finnish. In C. de Groot & H. Tommola (eds.), *Aspect Bound. A Voyage into the Realm of Germanic, Slavonic, and Finno-Ugrian Aspectology*. Foris, Dordrecht.
- Heinämäki, Orvokki 1994: Aspect as boundedness in Finnish. In Carl Bache, Hans Basböll & Carl-Erik Lindberg (eds.), *Tense, Aspect, and Action: Empirical and Theoretical Contributions to Language Typology. Empirical Approaches to Language Typology*, vol. 12. Mouton de Gruyter, Berlin.
- Helasvuo, Marja-Liisa 1988: Subjekteina ja objekteina toimivat nominilausekkeet puhutussa suomessa. Unpublished Licentiate's Thesis. Department of Finnish, University of Helsinki. Helsinki.
- Helasvuo, Marja-Liisa 1991: Syntaktiset rakenteet kertomuksen jäsennyksen osoittimina. *Virittäjä* 1/1991, 33–47.
- Helasvuo, Marja-Liisa 1992: Argument Structure in Obolo. Ms. Department of linguistics, University of California, Santa Barbara.
- Helasvuo, Marja-Liisa 1996a: Ollako vai eikö olla — eksistentiaalilauseen subjektin kohtalonkysymys. *Virittäjä* 3/1996: 340–356.
- Helasvuo, Marja-Liisa 1996b: A Discourse Perspective on the Grammaticization of the Partitive Case in Finnish. In Timo Haukioja, Marja-Liisa Helasvuo & Elise Kärkkäinen (eds.), *SKY 1996 Yearbook of the Linguistic Association of Finland*. Pp. 7–34. Helsinki.
- Helasvuo, Marja-Liisa 1997a: What can intonation tell us about constituency? Paper given at the Workshop on Constituency and Discourse. University of California, Santa Barbara.
- Helasvuo, Marja-Liisa 1997b: Functions of case-marking vs. non-marking in Finnish discourse. In Anne-Marie Simon-Vandenberg, Kristin Davidse & Dirk Noël (eds.), *Reconnecting language: Morphology and syntax in functional perspectives*. CILT 154. Pp. 213–225. Benjamins, Amsterdam.
- Helasvuo, Marja-Liisa (in press): Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In Couper-Kuhlen & Selting (eds.) (in press).
- Helasvuo, Marja-Liisa (forthcoming): Argument splits in Finnish grammar and discourse. In J. W. Du Bois, L. E. Kumpf & W. J. Ashby (eds.), *Preferred argument structure: Grammar as architecture for function*. John Benjamins, Amsterdam.
- Herring, Susan C. 1989: Verbless Presentation and the Discourse Basis of Ergativity. *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago.

- Himmelman, Nikolaus 1997: Projective constituent structures. Paper given at the Workshop on Constituency and Discourse. University of California, Santa Barbara.
- Hopper, Paul J. 1987: Emergent Grammar. *Proceedings of the thirteenth annual meeting, Berkeley Linguistic Society*. Pp. 139–157. Berkeley Linguistic Society, Berkeley.
- Hopper, Paul J. 1988: Emergent Grammar and the A Priori Grammar Postulate. In Deborah Tannen (ed.), *Linguistics in Context*. Ablex, Norwood, NJ.
- Hopper Paul J. & Sandra A. Thompson 1980: Transitivity in Grammar and in Discourse. *Language* 56: 251–299.
- Hopper, Paul J. & Sandra A. Thompson 1984: The Discourse Basis for Lexical Categories in Universal Grammar. *Language* 60: 703–752.
- Huomo, Tuomas 1994: Näkökulmia suomen ja viron sanajärjestyseroihin. (Perspectives on word order differences between Finnish and Estonian.) In *Lähivertailuja* 8. Suomalais-virolainen kontrastiivinen seminaari Hailuodossa 7.-9.5.1994. Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja 40. University of Oulu, Oulu.
- Huomo, Tuomas 1995: Ongella Elmeri huitaisi Anselmia ongella: Lauseaseman vaikutuksesta eräiden adverbiaalityyppien tulkintaan. *Virittäjä* 99: 45–70.
- Huomo, Tuomas 1996a: Domain shifts and the grammaticalization of case: a case study of the Finnish adessive. *Folia Linguistica Historica* XVII/1–2: 73–95.
- Huomo, Tuomas 1996b: Bound Spaces and the Semantic Interpretation of Existentials. *Linguistics* 34 (2), 295–328.
- Huomo, Tuomas 1996c: A Scoping Hierarchy of Locatives. *Cognitive Linguistics* 7: 3.
- Inaba, Nobufumi 2000: Genetiivin ja partitiivin datiiivinen käyttö vanhassa kirjasuomessa: sijanmerkinnän ja sijajärjestelmän suhteesta. (English Summary: Dative use of the genitive and partitive in Old Finnish: On the relation between case-marking and the case system). *Sananjalka* 42: 47–86.
- Itkonen, Erkki 1972: Über das Objekt in den finnisch-wolgaischen Sprachen. *Finnisch-Ugrische Forschungen* 39: 153–213.
- Itkonen, Terho 1979: Subject and Object Marking in Finnish: An Inverted Ergative System and an “Ideal” Ergative Sub-system. In Frans Plank (ed.), *Ergativity*. Academic Press, New York.
- Jakobson, Roman 1971 [1936]: Beitrag zur allgemeinen Kasuslehre. *Selected Writings* II. Pp. 23–71. Mouton, the Hague.
- Jakobson, Roman 1984 [1958]: Morphological Observations on Slavic Declension. (The Structure of Russian Case Forms.) [Morfologičeskie nabljudenija nad slavljańskim sklonenijem. (Sostav russkix padežnyx form.)] In L. R. Waugh & M. Halle (eds.), *Russian and Slavic Grammar: Studies 1931–1981*. *Janua Linguarum Series Major* 106. Mouton, the Hague.
- Jefferson, Gail 1990: List-Construction as a Task and Resource. In G. Psathas (ed.), *Interactional Competence*. Pp. 63–92. International Institute for Ethnomethodology and Conversation Analysis & University Press of America, Washington D.C.
- Kangasmaa-Minn, Eeva 1983: Tense, Aspect, and Aktionsart in Finno-Ugrian. In C. de Groot & H. Tommola (eds.), *Aspect Bound. A Voyage into the Realm of Germanic, Slavonic, and Finno-Ugrian Aspectology*. Foris, Dordrecht.
- Kärkkäinen, Elise 1996: Preferred Argument Structure and Subject Role in American English Conversational Discourse. *Journal of Pragmatics* 25: 675–701.

- Karlsson, Fred 1982: *Suomen kielen äänne- ja muotorakenne*. WSOY, Porvoo.
- Keenan, Elinor & Bambi Schieffelin 1976: Foregrounding referents: a reconsideration of left dislocation in discourse. *Berkeley Linguistic Society* 2: 240–257.
- Kelomäki, Tapani 1998: *Ekvatiivilause*. (The equational clause.) Finnish Literature Society, Helsinki.
- Koivisto, Vesa 1995: *Itämerensuomen refleksiivit*. (Reflexives in the Baltic-Finnic Languages.) Finnish Literature Society, Helsinki.
- Kumpf, Lorraine 1987: The Use of Pitch Phenomena in the Structuring of Stories. In Russell S. Tomlin (ed.), *Coherence and Grounding in Discourse*. Pp. 189–216. John Benjamins, Amsterdam.
- Kumpf, Lorraine 1992: Preferred Argument Structure in Second Language Discourse: A Preliminary Study. *Studies in Language* 16(2): 369–403.
- Laberge, Suzanne 1978: The changing distribution of indefinite pronouns in discourse. In R.W. Shuy & A. Shnukal (eds.), *Language Use and the Uses of Language*. Georgetown University Press, Washington.
- Ladd, D. Robert 1986: Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3: 311–340.
- Laitinen, Lea 1992: *Välttämättömyys ja persoona. Suomen murteiden ja nesessiivisten rakenteiden semantiikkaa ja kielioppia*. (English abstract: Necessity and Person: The Semantics and Grammar of Necessitative Structures in Finnish Dialects.) Finnish Literature Society, Helsinki.
- Laitinen, Lea 1995: Nollapersoona. (English summary: The Zero Person.) *Virittäjä* 99: 337–358.
- Laitinen, Lea 1996: Indexical site to be identified with: The case of the “Zero Person” in Finnish conversation. Paper given in the 4th International Pragmatics Conference, July 1996. Mexico City, Mexico.
- Lambrecht, Knud 1981: Topic, Antitopic, and Verb Agreement in Non-Standard French. *Pragmatics and Beyond*, Vol. II: 6. John Benjamins, Amsterdam.
- Lambrecht, Knud 1987: On the Status of SVO Sentences in French Discourse. In Russell Tomlin (ed.), *Coherence and Grounding in Discourse*. Pp. 217–261. John Benjamins, Amsterdam.
- Lambrecht, Knud 1988: Presentational Cleft Constructions in Spoken French. In John Haiman & Sandra A. Thompson (eds.), *Clause Combining in Grammar and Discourse*. Pp. 135–179. John Benjamins, Amsterdam.
- Langacker, Ronald W. 1987: *Foundations of Cognitive Grammar*. Vol. 1. Stanford University Press, Stanford.
- Langacker, Ronald W. 1995: Conceptual Grouping and Constituency in Cognitive Grammar. In Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm* 3. Pp. 149–172. Hanshin, Seoul.
- Langacker, Ronald 1997: Constituency, Dependency, and Conceptual Grouping. *Cognitive Linguistics* 8–1: 1–32.
- Larjavaara, Matti 1991: Aspektuaalisen objektin synty. *Virittäjä* 95: 372–407.
- Larsen, Thomas W. 1981: Functional Correlates of Ergativity in Aguacatec. In *Proceedings of the seventh annual meeting of the Berkeley Linguistics Society*. Pp. 136–153.

- Laury, Ritva 1982: Object Case Marking in Finnish. Unpublished M.A. Thesis. Department of Linguistics, California State University, Fresno.
- Laury, Ritva 1991: On the development of the definite article *se* in spoken Finnish. In Maria Vilkuna (ed.), *SKY 1991 Yearbook of the Linguistic Association of Finland*. Helsinki.
- Laury, Ritva 1997: *Demonstratives in Interaction: The emergence of a definite article in Finnish*. John Benjamins, Amsterdam.
- Lehmann, Christian 1988: On the function of agreement. In M. Barlow & Charles A. Ferguson (eds.), *Agreement in natural language: Approaches, theories, descriptions*. Pp. 55–65. Center for the Study of Language and Information, Stanford.
- Leino, Pentti 1990: Spatial Relations in Finnish: a Cognitive Perspective. In I. Almqvist, P.-E. Cederholm & J. Lainio (eds.), *Från Pohjolas pörten till kognitiv kontakt. Vänskrift till Erling Wande den 9 maj 1990*. Stockholm Studies in Finnish Language and Literature 6. Department of Finnish. Stockholm: Stockholm University. 117–152.
- Leino, Pentti 1991: *Lauseet ja tilanteet*. Finnish Literature Society, Helsinki.
- Leino, Pentti 1993: *Polysemia — kielen moniselitteisyys*. Kieli 7. Department of Finnish, University of Helsinki, Helsinki.
- Leino, Pentti, Marja-Liisa Helasvuo, Petri Lauerma, Urpo Nikanne & Tiina Onikki 1990: *Suomen kielen paikallissijat konseptuaalisessa semantiikassa*. Kieli 5. Publications of the Department of Finnish, University of Helsinki. Helsinki.
- Lerner, Gene 1991: On the Syntax of Sentences in Progress. *Language in Society*, 20 (3): 441–458.
- Lerner, Gene 1996: On the "semi-permeable" character of grammatical units in conversation: conditional entry into the turn space of another speaker. In Elinor Ochs, Emanuel Schegloff & Sandra A. Thompson (eds.), *Interaction and Grammar*. Pp. 238–276. Studies in Interactional Sociolinguistics 13. Cambridge University Press, Cambridge.
- Longacre, Robert 1987: The Semantics of the Storyline in East and West Africa. *Journal of Semantics* 5: 51–64.
- Lyons, John 1977: *Semantics*. Cambridge University Press, Cambridge.
- Mithun, Marianne 1991: Active/Agentive Case Marking and Its Motivations. *Language* 67: 510–546.
- Nakayama, Toshihide 1997: Discourse-Pragmatic Dynamism in Nuu-chah-nulth (Nootka) Morphosyntax. Unpublished Ph. D. Dissertation. Department of Linguistics, University of California, Santa Barbara.
- Nakayama, Toshihide & Kumiko Ichihashi-Nakayama 1994: Discourse perspective for core-oblique distinction in Japanese. *Santa Barbara Papers in Linguistics* 5: 158–177.
- Nemvalts, Peep 1994: Mittatilaustyökälu vai Prokrusteen vuode? Lingvistiikan terminologiaa pohtien. Paper given at the 1st Scandinavian Conference in Finnish Linguistics May 26–28, 1994. University of Stockholm, Stockholm.
- Nichols, Johanna 1983: On direct and oblique cases. In *Proceedings of the ninth annual meeting of the Berkeley Linguistics Society*. Pp. 170–192.
- Nichols, Johanna & Alan Timberlake 1991: Grammaticalization as Retextualization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to Grammaticalization*. Vol. 1. Pp. 129–146. John Benjamins, Amsterdam.
- Nunberg, Geoffrey 1993: Indexicality and deixis. *Linguistics and Philosophy* 16: 1–43.

- Ochs, Elinor 1988: *Culture and Language Development: Language Acquisition and Language Socialization in a Samoan Village*. Cambridge University Press, Cambridge.
- Ochs, Elinor, Emanuel A. Schegloff & Sandra A. Thompson (eds.) 1996: *Interaction and grammar*. Studies in Interactional Sociolinguistics 13. Cambridge University Press, Cambridge.
- Onikki, Tiina 1995: Metonymic links in semantic extension. Polysemic tendencies in locative case expressions designating postures and facial expressions in Finnish. Paper given at the International Cognitive Linguistics conference. July 1995. University of New Mexico, Albuquerque, NM.
- Onikki, Tiina 1997: Kieliopin ja leksikon rajalla. In Tapani Lehtinen & Lea Laitinen (eds.), *Kieliopillistuminen: Tapaustutkimuksia suomesta*. Kieli 12. Pp. 6-19. Department of Finnish, University of Helsinki.
- Ono, Tsuyoshi & Sandra A. Thompson 1994: Unattached NPs in English Conversation. In *Proceedings of the 20th annual meeting of the Berkeley Linguistics Society*.
- Ono, Tsuyoshi & Sandra A. Thompson 1995: What can Conversation tell us about Syntax? In Philip Davis (ed.), *Alternative Linguistics: Descriptive and Theoretical Modes*. Current Issues in Linguistic Theory, vol. 102. Pp. 213-271. John Benjamins, Amsterdam.
- Pajunen, Anneli & Ulla Palomäki 1985: *Tilastotietoja suomen kielen rakenteesta 2*. (English summary: Frequency analysis of spoken and written discourse in Finnish). Finnish Research Centre for Domestic Languages, Helsinki.
- Palander, Marjatta 1991: Puhe- ja kirjakielen sanajärjestyseroja. (English summary: Word order differences in spoken and written language.) *Virittäjä* 95: 235-254.
- Payne, John R. 1985: Negation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Vol. 1, 197-242. Cambridge University Press, Cambridge.
- Penttilä, Aarni 1963: *Suomen kielioppi*. Second, revised edition. WSOY, Porvoo.
- Pierrehumbert, Janet Breckenridge 1980: The Phonology and Phonetics of English Intonation. Ph.D. dissertation, MIT.
- Prince, Ellen F. 1981: Toward a Taxonomy of Given-New Information. In P. Cole (ed.), *Radical Pragmatics*. Academic Press, New York.
- Raevaara, Liisa 1989: No — vuoronalkuinen partikkeli. In Auli Hakulinen (ed.), *Suomalaisen keskustelun keinoja I*. Kieli 4. Department of Finnish, University of Helsinki.
- Sacks, Harvey 1992a: *Lectures on Conversation*. Vol. 1. Edited by Gail Jefferson. Basil Blackwell, Cambridge.
- Sacks, Harvey 1992b: *Lectures on Conversation*. Vol. 2. Edited by Gail Jefferson. Basil Blackwell, Cambridge.
- Sacks, Harvey; Schegloff, Emanuel & Jefferson, Gail 1974: A Simplest Systematics for Turn-Taking for Conversation. *Language* 50: 696-735.
- Savijärvi, Ilkka 1977: *Itämerensuomalaisten kielten kieltöverbi*. Osa 1: *Suomi*. (English summary: The negative verb in the Baltic-Finnic Languages. Part 1: Finnish.) Finnish Literature Society, Helsinki.
- Schegloff, Emanuel 1979: The Relevance of Repair to Syntax-for-conversation. In Talmy Givón (ed.), *Discourse and Syntax*. Syntax and Semantics. Vol. 12. Academic Press, New York.

- Schegloff, Emanuel 1987: Recycled Turn Beginnings: A precise repair mechanism in conversation's turn-taking organization. In Graham Button & John R. E. Lee (eds.) *Talk and Social Organisation*. Pp. 70–85. Multilingual Matters, Clevedon, England.
- Schegloff, Emanuel 1991: Conversation Analysis and Socially Shared Cognition. In Lauren Resnick, John Levine and Stephanie Teasley (eds.), *Perspectives on Socially Shared Cognition*. Pp. 150–171. American Psychological Association, Washington D. C.
- Schuetze-Coburn, Stephan 1992: Prosodic Phrase as A Prototype. Paper read in the Workshop on Prosody in Natural Speech, 5–12 August 1992, University of Pennsylvania.
- Schuetze-Coburn, Stephan; Marian Shapley & Elizabeth Weber 1991: Units of Intonation in Discourse: A Comparison of Acoustic and Auditory Analyses. *Language and Speech* 34: 207–234.
- Seppänen, Eeva-Leena 1995: Pronominit tämä, tuo, se ja hän viittaamassa keskustelun osallistajaan. Unpublished Licentiate's Thesis. Department of Finnish, University of Helsinki.
- Seppänen, Eeva-Leena 1996: Ways of referring to a knowing co-participant in Finnish conversation. In Timo Haukioja, Marja-Liisa Helasvuo & Elise Kärkkäinen (eds.), *SKY 1996 Yearbook of the Linguistic Association of Finland*. Pp. 135–176. Helsinki.
- Setälä, E. N. 1880: *Suomen kielen lauseoppi*. K. E. Holm, Helsinki.
- Setälä, E. N. 1908[1898]: *Suomen kielioppi*. Otava, Helsinki.
- Siewierska, Anna 1997: The Formal Realization of Case and Agreement Marking: A Functional Perspective. In Anne-Marie Simon-Vandenberg, Kristin Davidse & Dirk Noël (eds.), *Reconnecting Language: Morphology and Syntax in Functional Perspectives*. Current Issues in Linguistic Theory, vol. 154. Pp. 181–210. John Benjamins, Amsterdam.
- Silverstein, Michael 1976: Hierarchy of Features and Ergativity. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. Pp. 112–171. Linguistic Series No. 22, Australian Institute of Aboriginal Studies, Canberra. Humanities Press, New Jersey.
- Silverstein, Michael 1981: Case marking and the nature of language. *Australian Journal of Linguistics* 1: 227–246.
- Shore, Susanna 1988: On the so-called Finnish Passive. *Word* 39:151–176.
- Shore, Susanna 1992: Aspects of a Grammar of Finnish. Unpublished Ph.D. Thesis, Macquarie University, Sydney.
- Siro, Paavo 1964: *Suomen kielen lauseoppi*. Tietosanakirja, Helsinki.
- Sorjonen, Marja-Leena 1989: Vuoron alun merkityksestä. In Auli Hakulinen (ed.), *Suomalaisen keskustelun keinoja* 1. Kieli 4. Pp. 129–146. Department of Finnish, University of Helsinki.
- Sorjonen, Marja-Leena 1996: On Repeats and Responses in Finnish Conversations. In Elinor Ochs, Emanuel A. Schegloff & Sandra A. Thompson (eds.), *Interaction and Grammar*. Studies in Interactional Sociolinguistics 13. Pp. 277–327. Cambridge University Press, Cambridge.
- Sorjonen, Marja-Leena 1997: Recipient Activities: Particles nii(n) and joo as Responses in Finnish Conversations. Unpublished dissertation. Department of Applied Linguistics, University of California, Los Angeles.
- Tainio, Liisa 1993: Kannanoton tulkinta keskustelussa. Unpublished licentiate's thesis. Department of Finnish, University of Helsinki.

- Tao, Hongyin 1992: NP Intonation Units and Referent Identification. In *Proceedings of the eighteenth annual meeting of the Berkeley Linguistics Society*.
- Tao, Hongyin 1996: *Units in Mandarin Conversation*. John Benjamins, Amsterdam.
- Terken, Jacques & Julia Hirschberg 1994: Deaccentuation of words representing 'given' information: Effects of persistence of grammatical function and surface position. *Language and Speech* 37(2): 125–145.
- Thompson, Sandra A. 1997: Discourse Motivations for the Core-Oblique Distinction as a Language Universal. In Akio Kamio (ed.), *Directions in Functional Linguistics*. John Benjamins, Amsterdam.
- Tiainen, Outi 1997: Subjektin rooli diskurssissa. Unpublished M.A. Thesis. Department of Finnish, University of Helsinki.
- Toivainen, Jorma 1986: Suomen subjektin piirteitä. (English summary: Some features of the Finnish subject.) *Sananjalka* 28: 31–46.
- Tuomikoski, Risto 1971: Persoona, tekijä ja henkilö. *Virittäjä* 75: 146–152.
- Vilkuna, Maria 1989: *Free Word Order in Finnish*. Finnish Literature Society, Helsinki.
- Vilkuna, Maria 1996: *Suomen lauseopin perusteet*. Edita, Helsinki.
- Wähämäki, Börje 1975: Om finskans existentialsatser och därtill hörande problem. Unpublished Licentiate's Thesis. Department of Finnish, Åbo Akademi, Åbo.
- Wähämäki, Börje 1984: *Existence and identity: A study of the semantics and syntax of existential sentences in Finnish*. Publications of the Research Institute of the Åbo Akademi Foundation. Vol. 99. Åbo.
- Yli-Vakkuri, Valma 1986: *Suomen kielipiillisten muotojen toissijainen käyttö*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 28. University of Turku, Turku.

Appendix

Glossing and transcription conventions

Glossing

The morphemes have been indicated with dashes; in the case of fused morphemes, there is a plus sign in the glossing. The following form classes have been treated as unmarked and are not indicated in the glossing: nominative case and singular form in the nominal markings, present tense and active voice in the verbal markings.

Nominal markings

ABL	ablative
ACC	accusative
ADE	adessive
ALL	allative
ELA	elative
ESS	essive
GEN	genitive
ILL	illative
INE	inessive
PTV	partitive
TRA	translative
COM	comitative
PL	plural
COMP	comparative form (adjectives only)
SUP	superlative form (adjectives only)
1SG	first person singular pronoun (likewise 2nd and 3rd)
1PL	first person plural pronoun

Verbal markings

1SG	first person singular person marking (likewise 2nd and 3rd)
1PL	first person plural person marking
PAST	past tense
COND	conditional mood
IMP	imperative mood
INF	infinitive
PCP	participle
NEG	negation verb
PSS	passive

Other

Q	question clitic
CLT	other clitics
PX	possessive suffix

Transcription conventions*Transitional continuity*

,	continuing intonation (slightly falling)
.	terminal intonation (falling)
?	appeal (slightly rising)
—	truncated intonation

Accent and lengthening

^	primary accent
=	lengthening of the preceding sound

Quality

⟨F F⟩	forte: loud
⟨P P⟩	piano: soft
⟨HI HI⟩	higher pitch level
⟨LO LO⟩	lower pitch level
⟨A A⟩	allegro: rapid speech
⟨L L⟩	lento: slow speech

⟨%%⟩ creaky voice
⟨@ @⟩ laughing voice

Vocal noises

@ laughter
(THROAT) someone clearing their throat
(H) inbreath

Other

⟨X X⟩ uncertain hearing
...() pause (length indicated in parantheses)

Name index

A

Andrews 33
Ashby 12, 14, 53, 67, 93, 108, 113, 126,
127
Auer 16

B

Bentivoglio 12, 14, 53, 93, 108
Blake 13, 64
Brown 135
Bruun 10
Bybee 3, 7, 62

C

Campbell 35
Chafe 2, 3, 4, 12, 13, 16, 21, 42, 81, 90,
95, 96, 126, 127, 133, 140, 147, 148
Comrie 13
Couper-Kuhlen 3, 16, 21, 140
Croft 33, 64, 112, 134, 135, 136
Cruttenden 16, 133
Crystal 134
Cumming 3

D

Dahl 78
Davis 3, 21
Dixon 13
Du Bois 2, 3, 6, 8, 12, 13, 14, 15, 16, 25,
34, 45, 53, 56, 66, 81, 86, 87, 91,
93, 109, 110, 114, 133
Duranti 3
Durie 13, 14, 23, 34, 45, 54, 60, 91, 108,
110, 111, 112, 113

E

Erringer 136

F

Foley 22, 35, 57
Ford 3, 21, 25, 144
Forsberg 28
Fox 3, 21, 144, 148
Franck 69

G

Geluykens 3, 108, 110, 126, 127, 129,
130
Givón 2, 22, 34, 53, 55, 64, 129, 134
Goffman 74
Goodwin, M. H. 3
Goodwin, C. 2, 3, 20,

H

Haiman 134
Hakanen 63
Häkkinen 51
Hakulinen A. 7, 31, 37, 76, 77, 89, 100
Hakulinen L. 61, 65
Halliday 16, 24, 57, 108, 133
Harris 35
Hayashi 21
Heinämäki 36, 47, 101
Heine 1, 6, 35, 97
Helasvuo 23, 35, 37, 44, 51, 52, 54, 55,
58, 60, 97, 101, 108, 111, 112, 123,
136, 143, 146, 147
Herring 12, 23, 24, 53, 93, 109, 110, 111,
113, 127
Himmelmann 137
Hirschberg 12
Hopper 1, 2, 3, 25, 52, 53, 105, 131, 147,
148
Huumo 61, 62, 76, 77, 91, 96, 100

I

Ichihashi-Nakayama 53, 93
Inaba 61
Itkonen E. 49
Itkonen T. 36

J

Jakobson 35
Jaspersen 21, 148
Jefferson 3, 20

K

Kang 146
Kangasmaa-Minn 112
Kärkkäinen 14, 45, 53, 93, 146
Karlsson 7, 37, 39, 42, 57, 89, 100
Keenan 126
Koivisto 65
Kumpf 14, 148

L

Laberge 67
Ladd 133
Laitinen 68, 69, 72
Lambrecht 12, 21, 45, 81, 108, 113, 126
Langacker 4, 6, 8, 135, 146
Larjavaara 101
Larsen 12,
Laury 52, 137, 138
Lehmann 64, 72, 74, 105
Leino 36, 38, 61, 101
Lerner 3, 20, 122
Lindén 78
Longacre 111
Lyons 93, 97

M

Mallinson 13, 64
Mithun 13, 34
Mulac 146

N

Nakayama 53, 83, 93, 153
Nemvalts 38
Nichols 8, 35

Nunberg 67, 74

O

Ochs 3, 21, 45
Onikki 61
Ono 2, 3, 4, 8, 9, 23, 108, 111, 112, 113,
117, 123, 125
Östman 146

P

Pajunen 43
Palander 44, 45
Palomäki 43
Payne J. 78
Penttilä 37
Pierrehumbert 16, 133
Prince 12

R

Raevaara 32
Räsänen 39

S

Sacks 20, 110, 122
Savijärvi 78
Schegloff 3, 4, 8, 21, 110
Scheibman 3, 146
Schieffelin 126
Schuetze-Coburn 25
Selting 3, 21, 140
Seppänen 30
Setälä 37, 49
Shore 36, 37, 50, 65
Siewierska 34
Silverstein 13, 14, 34, 94, 95, 96, 152
Siro 36, 123
Sorjonen 3, 30, 31, 121, 128, 143

T

Tainio 93
Tao 3, 21, 22, 23, 25, 33, 83, 108, 110,
112, 126, 153
Terken 12
Thompson 2, 3, 4, 8, 9, 13, 21, 23, 25,
52, 53, 60, 92, 105, 108, 111, 112,

113, 117, 123, 125, 126, 131, 144,
146, 147

Tiainen 66

Timberlake 8

Toivainen 101

Tuomikoski 65

V

Van Valin 22, 35, 57,

Vilkuna 7, 38, 46, 50, 63, 75, 76, 78, 89,

95, 100, 148

W

Wähämäki 97, 99, 102

Y

Yli-Vakkuri 28

Subject index

A

activation cost 12
 new mentions 12, 54–55, 90–91,
 99–101, 109–111, 147
aspectual boundedness 47, 101

C

clause, see grammatical units
constituency 135–140, 146, 148
case marking
 core cases 35–40, 56–57
 oblique cases 35–40, 57

D

discourse referentiality 13
 tracking NPs 13, 54–55, 60–61, 91–94

E

E-NP, see existential construction
emergent grammar, see grammar
existential construction 7, 61–63, 97–101
 E-NP 7–8, 97–101, 103

F

free NPs, see grammatical units

G

grammar 3–4, 8–9
 emergent grammar 1, 151–154
grammatical units 1, 2, 3, 108–113
 clause 1, 3, 21–33, 133–135
 free NPs 3, 23–24, 29–31, 105–131
 particle chains 31–33
grammaticization 1–2, 6, 8, 35–36,
 61–63, 94–96, 101, 103

H

human referents 54–55, 79–81, 86–87,
 109–110

I

intonation units 16–17, 133

M

marking of personal pronouns 34, 43,
 45, 46, 49–50, 94–96, 103

N

new mentions, see activation cost

O

object role 14–15, 139–140, 146–148
oblique cases, see case marking

P

particle chains, see grammatical units
passive 14–15, 66–67, 72
preferred argument structure 86–87,
 109–110
prototypes 4

S

schemas 4, 6, 7, 8
sentence 2, 19–20
sequential organization 6
split patterns in alignment systems
 34–35, 94–96, 103
subject role 6–8, 13–14, 85–103
syntactic units, see grammatical units

T

tracking NPs, see discourse
 referentiality
transitional continuity 16–17
transitivity 52–53, 147–148

Z

zero person 69, 72

In the series STUDIES IN DISCOURSE AND GRAMMAR (SiDaG) the following titles have been published:

1. GELUYKENS, Ronald: *From Discourse Process to Grammatical Construction: On Left-Dislocation in English*. Amsterdam/Philadelphia, 1992.
2. IWASAKI, Shoichi: *Subjectivity in Grammar and Discourse: Theoretical Considerations and a Case Study of Japanese Spoken Discourse*. Amsterdam/Philadelphia, 1993.
3. WEBER, Elizabeth G.: *Varieties of Questions in English Conversation*. Amsterdam/Philadelphia, 1993.
4. DOWNING, Pamela: *Numerical Classifier Systems: The Case of Japanese*. Amsterdam/Philadelphia, 1996.
5. TAO, Hongyin: *Units in Mandarin Conversation: Prosody, Discourse, and Grammar*. Amsterdam/Philadelphia, 1996.
6. DORGELOH, Heidrun: *Inversion in Modern English: Form and function*. Amsterdam/Philadelphia, 1997.
7. LAURY, Ritva: *Demonstratives in Interaction. The emergence of a definite article in Finnish*. Amsterdam/Philadelphia, 1997.
8. MORI, Junko: *Negotiating Agreement and Disagreement in Japanese. Connective expressions and turn construction*. Amsterdam/Philadelphia, 1999.
9. HELASVUO, Marja-Liisa: *When Discourse becomes Syntax. Noun phrases and clauses as emergent syntactic units in Finnish conversational discourse*. Amsterdam/Philadelphia, 2001.
10. SELTING, Margret and Elizabeth COUPER-KUHLEN (eds.): *Studies in Interactional Linguistics*. n.y.p.