

Objects and information structure

Mary Dalrymple (mary.dalrymple@ling-phil.ox.ac.uk)
Centre for Linguistics and Philology
Oxford University
Walton Street
Oxford OX1 2HG UK
telephone: +44 1865 280 403 fax: +44 1865 280 412

Irina Nikolaeva (irina_a_nikolaeva@yahoo.com)
Department of Linguistics
School of Oriental and African Studies
University of London
Thornhaugh Street, Russell Square
London WC1H 0XG
telephone: +44 (0)20 7898 4250

DRAFT: Do not quote or cite

March 2, 2010

List of Abbreviations

We have regularised some of the glosses and the transcription of the original sources. We use the following abbreviations:

Abs	absolutive	Fem	feminine	Obl	oblique agreement
Acc	accusative	Foc	focus	P	potential
All	allative	Fut	future	Part	participle
AN	action nominal	Gen	genitive	Pas	passive
Art	article	Hab	habitual	Perf	perfective
Asp	aspect	Hon	honorific	Pl	plural
Attr	attributive	Impf	imperfective	Poss	possessive
Aug	augmentative	Imprs	impersonal	Postess	postessive
Aux	auxiliary	Imp	imperative	Prep	preposition
Caus	causative	Indef	indefinite	Pres	present
Cl	classifier	Indic	indicative	Prog	progressive
Com	comitative	Inf	infinitive	Proh	prohibitive
Conv	converb	Instr	instrumental	Pron	pronominal
Cop	copula	Interj	interjection	Purp	purposive
C	completive	Intr	intransitive	Rel	relative form
Dat	dative	Irr	irrealis	Rep	reportative
Decl	declarative	Loc	locative	Sg	singular
Det	determiner	Masc	masculine	Spec	specific
Dir	directive	Mod	modal	Subj	subject agreement
Dur	durative	Neg	negation	Sub	subordinate
Du	dual	NMLZ	nominalisation	Supess	superessive
Erg	ergative	Nom	nominative	Top	topic
Expl	expletive	Obj	object marker or	Tr	transitive
Ez	ezafe		object agreement		

Preface and Acknowledgements

Our work on this topic began in connection with our participation in the AHRC project “Noun Phrase Agreement and Coordination” (AHRC grant MRG-AN10939/APN17606, 2004-2006), a joint project which involved, besides us, Louisa Sadler, Ryo Ootoguro, and Aline Villavincencio at the University of Essex. We are grateful to Louisa, Ryo, and Aline for helpful comments and discussion as our ideas developed over the course of the project. Earlier versions of the work were presented at the 2005 International Lexical-Functional Grammar Conference (LFG05) in Bergen, the 2005 Linguistic Association of Great Britain conference in Cambridge, the 2005 Colloque de Syntaxe et Sémantique à Paris (CSSP 2005), the University of Leipzig in December 2007, the University of Manchester in April 2008, and the workshop “Arctic Languages: Syntax, Morphology, Lexicon”, University of Tromsø in September 2009. An early version of our work appeared as Dalrymple and Nikolaeva (2005), though it will be clear to readers of that paper that our ideas have evolved considerably since its publication. For very helpful comments and discussion of these earlier versions, we are grateful to Farrell Ackerman, Ash Asudeh, David Beck, Balthasar Bickel, Miriam Butt, Greville Corbett, David Cram, Annahita Farudi, Ron Kaplan, Tracy Holloway King, Aditi Lahiri, Olivia Lam, Louise Mycock, Tara Mohanan, Louisa Sadler, Devyani Sharma, Nigel Vincent, Yogendra Yadava, and Annie Zaenen. Farrell Ackerman, Louise Mycock, and Louisa Sadler provided very helpful comments on drafts of this book as it took shape. We are grateful to our language consultants for their generous assistance: Ghazaleh Kad and Shamsi Saber for Persian, Dolgor Guntsetseg and Tuya Bavuu for Khalkha Mongolian, Galina Koreneva and Anna Lamdo for Tundra Nenets, and Devyani Sharma and Tara Mohanan for Hindi. Finally, we thank our husbands, Ken Kahn and Simon Carne, for their love and support during the writing of this book.

Contents

List of Abbreviations	i
Preface	iii
1 Introduction	1
1.1 The phenomenon	1
1.2 Previous work	2
1.2.1 Marking as distinguishing arguments	3
1.2.2 Marking as coding features	6
1.2.3 DOM in transformational syntax	8
1.3 Criteria for marking	10
1.4 Our proposal	14
1.4.1 Information structure	14
1.4.2 Syntax	17
1.4.3 Diachrony and grammaticalisation	18
1.4.4 Limits of our analysis	19
1.5 Structure of the book	20
2 Syntactic assumptions	23
2.1 Grammatical functions	24
2.1.1 Grammatical function diagnostics	24
2.1.2 Grammatical functions and semantic roles	26
2.2 Objects	27
2.3 Levels of syntactic representation	29
2.4 Describing linguistic structures	33
2.5 Grammatical agreement	36
2.6 Agreement and pronominal incorporation	37
2.7 Casemarking	40
2.8 Nonsyntactic criteria for casemarking patterns	43
2.9 Conclusion	45

3	Information structure in grammar	47
3.1	The content of information structure	47
3.2	Information structure roles	49
3.2.1	Focus	49
3.2.2	Topic	50
3.2.3	Topicality and topic-worthiness	52
3.2.4	Secondary topic	55
3.3	Conclusion	59
4	Syntax and information structure	61
4.1	Alternative views of information structure	61
4.1.1	Tree-based representations	61
4.1.2	Early work in LFG	64
4.2	Information structure and its role in grammar	68
4.2.1	Content of information structure: The LFG view	68
4.2.2	Linguistic encoding of information structure relations	69
4.2.3	Information structure in relation to semantics	71
4.3	Our architecture	74
4.3.1	Glue and the syntax-semantics interface	74
4.3.2	Information structure features and relations	77
4.3.3	Levels and equations	81
4.3.4	A short text	88
4.3.5	Information structure and its place in grammar	91
4.4	Conclusion	93
5	Topicality and grammatical marking	95
5.1	Topical marking for different grammatical functions	95
5.2	Grammatical encoding of topical subjects	100
5.3	Subjects and topichood	103
5.4	Conclusion	107
6	Topical marking of nonsubjects	109
6.1	Casemarking of topical nonsubjects	109
6.1.1	Persian	109
6.1.2	Tariana	115
6.1.3	Topical nonsubject casemarking	118
6.2	Agreement with topical nonsubjects	119
6.2.1	Itelmen	119
6.2.2	Tabassaran	122
6.2.3	Topical nonsubject agreement	125
6.3	Conclusion	127

7	Topicality and DOM	129
7.1	Objects as grammaticalised secondary topics	129
7.2	Agreement with topical objects: Tundra Nenets	131
7.3	Casemarking of topical objects	137
7.3.1	Tigre	137
7.3.2	Dolakha Newar	141
7.4	Conclusion	143
8	Primary and secondary objecthood and DOM	145
8.1	Grammatical marking and grammatical function	145
8.2	Object agreement and grammatical function: Ostyak	146
8.3	Object casemarking and grammatical function	156
8.3.1	Mongolian	156
8.3.2	Chatino	160
8.3.3	Hindi	164
8.4	Objects and markedness	170
8.5	Conclusion	173
9	Multiple objects and grammatical alignment	175
9.1	On the typology of multitransitive constructions	175
9.2	No ditransitive construction	178
9.2.1	Nenets	178
9.2.2	Ostyak	179
9.2.3	Mongolian	182
9.3	Ditransitive constructions	183
9.3.1	Chatino	184
9.3.2	Hindi	185
9.4	Other multitransitive constructions	186
9.4.1	Dolakha Newar	187
9.4.2	Tigre	191
9.4.3	Applicatives: Upper Necaxa Totonac	194
9.5	Topicality and goals	198
9.6	Conclusion	200
10	Grammaticalisation	201
10.1	Case studies	201
10.1.1	Uralic	202
10.1.2	Persian and the Iranian languages	209
10.1.3	Hindi and the Indo-Aryan languages	210
10.2	Paths of grammaticalisation	214
10.2.1	Spreading of DOM	214
10.2.2	Narrowing of DOM	218

10.3 Towards a typology	221
10.4 Conclusion	224
11 Conclusion	225

1

Introduction

1.1 The phenomenon

Many languages exhibit non-uniform grammatical marking targeting objects. Variations can occur within one and the same language with objects of one and the same verb. For example, in Turkish (Altaic¹) the object of the same verb either takes the accusative suffix or remains unmarked:²

- (1) a. Ali bir kitab-i aldı
Ali one book-**Acc** buy.Past.3Sg
'Ali bought a certain book.'
- b. Ali bir kitap aldı
Ali one book buy.Past.3Sg
'Ali bought a book.' (Enç 1991:5)

In Palauan (Austronesian), the object of the same verb either does or does not trigger agreement on that verb:

- (2) a. Te-'illebed-ii a bilis a rengalek
Subj.3Pl-Perf.hit-**Obj**.3Sg the dog the children
'The kids hit the dog.'
- b. Te-'illebed a bilis a rengalek
Subj.3Pl-Perf.hit the dog the children
'The kids hit a dog/the dogs/some dog(s).' (Woolford 2000:5)

¹Language families are provided according to the classifications in Ethnologue (Grimes 1999).

²References are provided for examples that do not come from our own fieldwork.

Such patterns are widely known under the rubric of **differential object marking** or DOM (a term introduced by Bossong 1985).

We understand DOM as covering both agreement and casemarking (case or adpositional marking on the object). Though we recognise that agreement and casemarking differ both historically and synchronically, as noted by Comrie (1979) and Croft (1988:167–168), among many others, we believe that they share commonalities in DOM, and we will use the cover term **(grammatical) marking** to refer to them. This approach is in line with Nichols (1986), who analyses agreement and casemarking as alternative strategies for encoding the relation between the head and a dependent, as well as some generative literature, where case and agreement are inherently linked.

The aim of this book is to provide a new view of DOM which encompasses syntactic, semantic, and information-structural differences between marked and unmarked objects. We will make the following claims:

- Marked objects are associated with the information-structural role of **topic**. The association is synchronic in some languages, and may be historical in others. Where the direct connection between marked objects and topicality has been lost through grammaticalisation, marked objects in some languages become associated with **semantic features** typical of topics (animacy, definiteness, specificity).
- In some languages, marked and unmarked objects display an identical behavioural profile and can be assigned to the same grammatical function. Other languages distinguish syntactically between marked and unmarked objects: marked objects are **primary objects**, while unmarked objects are **secondary objects**. This reflects the tendency for topical arguments to appear high on the grammatical function hierarchy, and nontopical arguments to appear lower.

We begin our discussion with a review of previous work; we then present the essential aspects of our claims, and conclude this chapter with an overview of the book.

1.2 Previous work

DOM has been studied from a formal, generative perspective as well as a functional-typological perspective, and has been discussed and analysed in detail by Lazard (1984), Bossong (1985, 1991), de Hoop (1992), Aissen (2003a,b), Næss (2004), and de Swart (2007), among many others. Many of these analyses concentrate either on differential object agreement or on differential object casemarking, including both case and adpositional marking.

1.2.1 *Marking as distinguishing arguments*

Analyses of grammatical marking (and in particular casemarking) have often appealed to two types of functional motivation, referred to as **coding/indexing** and **discriminatory/disambiguating/distinguishing** (Comrie 1979, 1989, de Hoop and Narasimhan 2005, de Hoop and Malchukov 2007, Næss 2007, Malchukov 2008, and others).

Discriminatory or disambiguating casemarking serves to distinguish between different categories: for example, between the two arguments (the subject and the object) of a transitive clause. It encodes the **relation** between two arguments rather than the properties of an individual argument. The discriminatory function of casemarking has been argued to provide a functional motivation for the fact that in many languages casemarking is missing on the single argument of intransitive verbs and on one of the two arguments of transitive verbs. Silverstein (1976) and Comrie (1977) argue that since the basic purpose of formal marking on core arguments is to distinguish the subject from the object, the need to overtly mark the object is greater in some cases than in others because an object with subject-like semantic properties — for example, an animate, specific, or definite object — is more likely to be confused with the subject. Therefore, objects whose semantic features are typical of subjects are more likely to be overtly marked. This approach relies on the concept of the transitive prototype, in which the object is prototypically inanimate, indefinite, and/or nonspecific (Comrie 1989), and maintains that the function of DOM is to signal deviation from the prototype. It also stands in conformance with the widespread functional view that infrequent (and therefore functionally marked) categories receive more formal marking, whereas frequent (and therefore functionally unmarked) categories tend to remain formally unmarked; the explanation for this is based on economy considerations (Haiman and Thompson 1985, Du Bois 1987) and the iconic relation between functional and formal markedness. On this view, DOM is essentially iconic: formal marking on objects reflects their status as atypical or infrequent objects, and thus their functional markedness.

The idea that marking serves to distinguish or differentiate between arguments of a predicate has been pursued in depth in the influential work of Aissen (2003a,b). In her approach, languages may appeal to different factors in DOM, but in all cases the resulting patterns reflect the tension between two functional principles: (i) iconicity between functional and formal markedness of objects, and (ii) economy, the pressure to avoid excessive marking. Following much work in functional typology (Silverstein 1976, Givón 1976, Comrie 1977, 1979, 1989, Du Bois 1987, Dixon 1994), Aissen claims that there is a prototypical association involving grammatical functions and features such as animacy, humanness, definiteness and specificity/referentiality. Subjects are

prototypically high in these features, while objects are low. In other words, properties that are unmarked for subjects are marked for objects, a relation known as **markedness reversal**. On this view, unmarked subjects are animate, human, definite and specific, while marked subjects are inanimate and/or non-specific. For objects, the opposite markedness patterns are at work: inanimate and indefinite/nonspecific objects are unmarked, while nonhuman definite animate objects are more marked, and human objects are most highly marked.

Aissen formalises these correlations as Optimality Theoretic constraints. Referential properties of animacy, humanness, definiteness and specificity are organised into two Prominence Scales, the Animacy Scale and the Definiteness Scale.

- (3) Animacy Scale: Human > Animate > Inanimate
 Definiteness Scale: Personal pronoun > Proper name > Definite NP >
 Indefinite specific NP > Nonspecific NP

Humans are located higher on the Animacy Scale than nonhuman animates, which in turn are higher than inanimates, and so on. In addition, Aissen introduces a binary Relational Scale, where the subject outranks the object, as well as several harmonic (or markedness) hierarchies representing the relation between the Prominence Scales and the Relational Scale. The harmonic alignment constraints for animacy and definiteness features are displayed in (4) and (5), respectively.

- (4) *Su/Inan ≫ *Su/Anim ≫ *Su/Hum
 *Oj/Hum ≫ *Oj/Anim ≫ *Oj/Inan
- (5) *Su/NSpec ≫ *Su/Spec ≫ *Su/Def ≫ *Su/PN ≫ *Su/Pro
 *Oj/Pro ≫ *Oj/PN ≫ *Oj/Def ≫ *Oj/Spec ≫ *Oj/NSpec

The most highly ranked constraints in (4) penalise inanimate subjects and human objects; the constraints ranked one step lower penalise animate nonhuman subjects and animate nonhuman objects, and so on. The definiteness alignment constraints in (5) work similarly.

These hierarchies predict the most and least marked patterns of subject and object marking across languages. Constraints higher on the hierarchy incur more costly violations than constraints lower on the hierarchy. This means that if an object at a certain point in the hierarchy is overtly marked, then any object that is higher on the relevant scale will also be overtly marked. DOM arises when some but not all objects are marked; this is implemented by correlating harmonic constraint hierarchies with the constraint *STRUC_C, motivated by

the needs of economy, which penalises a value for the morphological category CASE below a certain point on the hierarchy.

The position of *STRUC_C in the hierarchy leads to different patterns of object marking across languages. If *STRUC_C dominates all the constraints on both scales, then marking is banned for all objects. If *STRUC_C is ranked at the lowest point on the hierarchy, all objects receive grammatical marking. Such languages do not display DOM. In languages with DOM, object marking can be based either on the Animacy Scale or on the Definiteness Scale. For example, Aissen (2003b) shows that in Catalan (Romance) only personal pronoun objects are casemarked. This is captured in an Optimality Theoretic account by positioning *STRUC_C lower in the Definiteness Scale than the constraint penalising pronominal objects. Similarly, if *STRUC_C is ranked lower than the top-ranked constraint *Oj/Hum in the Animacy Scale, casemarking is penalised for all objects except the most highly ranked human objects. According to Aissen (2003b:456), such languages are difficult to find, although there are many languages where marking is penalised for all objects except animates (including humans): for example, Sinhalese (Indo-Aryan). Further demotion of the economy constraint *STRUC_C yields other patterns of object marking. In Pitjantjatjara (Pama-Nyungan), pronominal and proper name objects are marked, while other objects, including definite objects, are unmarked; this is treated by positioning *STRUC_C below the constraints penalising marking on pronominal and proper name objects. In Hebrew (Semitic), only definite objects require the object marker *et*, while indefinite objects are always unmarked; this means that the economy constraint is ranked lower than the constraint penalising definite objects. Simultaneous reference to multiple features involves more complicated multidimensional crossing of the scales, but the basic principle remains the same.

Aissen's pioneering proposals have inspired much subsequent work and discussion, including an exploration of patterns that do not fit neatly into Aissen's cross-linguistic picture. For example, Yang and van Bergen (2007) argue that in Mandarin Chinese (Sino-Tibetan), objects that are obligatorily marked in the *ba*-construction are either highly prominent in terms of animacy or, surprisingly, **low** in prominence in terms of definiteness; for a small category of objects in-between, casemarking is optional. Importantly, the effect of animacy and definiteness is only noticeable on scrambled objects; Yang and van Bergen propose that the syntactic position of the object introduces an additional dimension into the prominence hierarchy which can influence marking. Other works following on from Aissen's work, including Morimoto (2002), propose various readjustments of the original constraint hierarchy, but do not question the general spirit of the prototype deviation approach to DOM.

1.2.2 Marking as coding features

The coding/indexing perspective on marking differs from the discriminatory perspective in that marking is taken to signal specific semantic and/or pragmatic properties of the relevant argument, rather than a particular relation between one argument and another. In fact, de Hoop and Narasimhan (2005) claim that a purely discriminatory function for casemarking is cross-linguistically rare, and that in most instances casemarking serves to mark some aspect of the interpretive content of a phrase. This is particularly clear for obliques, where case can bear an important semantic load in signalling the meaning contribution of the casemarked phrase. In the analysis of DOM, the coding/indexing perspective assumes that the presence of marking is connected to the presence of certain properties of the object. This view goes back to Hopper and Thompson's (1980) classic study, in which DOM is taken to be one of the signals indicating high transitivity, rather than deviation from a transitive prototype.

Næss (2004, 2006, 2007) argues that Aissen's approach contradicts these established notions of transitivity, and proposes that the prototypical transitive clause is one in which the two participants are maximally semantically distinct. Her definition of semantic distinctness includes several parameters, but the parameter that is especially relevant for cross-linguistic patterns of DOM is **affectedness**, understood roughly as involvement in the verbal event and change of state of the participant as a direct result of this event. On this view, prototypical objects are those that are highly affected by the transitive event. According to Næss, then, formally marked objects are not functionally marked; instead, they are "prototypical" highly affected and individuated objects, which tend to receive more grammatical marking than "non-prototypical", less affected objects.

Other researchers have also appealed to affectedness as a factor in analyses of DOM: for example, Çetinoğlu and Butt (2008) discuss the role of affectedness in object casemarking in Turkish. However, the relevance of affectedness for DOM has been questioned by de Hoop (2008), who shows that in many cases object marking is present in sentences in which the object participant is not affected by the verbal event. For instance, in Mandarin Chinese the object marker *ba*, often treated as an instance of DOM, is required on objects of the verbs 'forget' and 'lose', although the forgotten or lost thing is not usually affected by the forgetting or losing event. Næss (2004) claims that definite objects are more affected than indefinite ones because the event affects the whole rather than a part (cf. *I drank the milk* and *I drank some milk*). However, in many cases it is difficult to argue that definite or animate objects are more affected than indefinite or inanimate ones, if affectedness is understood in terms of a change of state. De Hoop cites the following Hindi (Indo-Aryan) examples from Mohanan (1994), involving the accusative marker *ko*:

- (6) a. Ilaa-ne haar uṭhaayaa
 Ila-Erg necklace lifted
 ‘Ila lifted a necklace.’
- b. Ilaa-ne haar-ko uṭhaayaa
 Ila-Erg necklace-Acc lifted
 ‘Ila lifted the necklace.’

(de Hoop 2008, citing Mohanan 1994:80)

According to de Hoop, there is no reason to think that the necklace that is picked up in (6b) is more affected than the necklace that is picked up in (6a), even though the former is definite and the latter is indefinite.

In fact, de Hoop and Narasimhan (2005), Næss (2007), and Malchukov (2008) note that in the case of DOM, disambiguating and indexing approaches lead to roughly the same result (though they make different predictions with respect to differential subject marking). Affectedness normally presupposes high individuation of the referent, while individuation in its turn is associated with definiteness. Animacy may also be relevant for affectedness because the effect of an action on an animate entity is more salient for human perception than the effect on an inanimate entity and is more likely to arouse empathy. Thus, affectedness is ultimately “operationalised” in terms of the same features of prominence as are relevant on the disambiguating/discriminatory marking perspective: when a language decides what kinds of objects are affected enough to be marked, it can choose on the basis of more easily measurable properties such as animacy and definiteness.

De Swart (2006, 2007) proposes an analysis which, in a sense, combines the discriminatory and coding approaches. His model is based on the idea that the speaker takes the perspective of the addressee in order to ensure **recoverability** of the intended interpretation. In some instances this implies that the speaker chooses to mark an object, rather than leaving it unmarked, when he/she wishes to emphasise a certain feature of the object: for instance, definiteness in Hindi. If the speaker intends to highlight definiteness in order to ensure that the addressee will interpret the object as definite, accusative *ko* appears. If the speaker does not want to force a definite interpretation of the object, casemarking is omitted. Thus, marking on the object is the result of a principled decision on the part of the speaker and has the coding function.

Recoverability presupposes “semantic distinctness” between two arguments (Næss 2004). This explains why casemarking on objects can be influenced by the semantic properties of the subject and vice versa. However, de Swart does not account for these patterns in terms of transitivity and, unlike Aissen, does not appeal to prototypical features of subjects and objects. He illustrates his

analysis with data from languages which seem to best support the disambiguating/discriminatory view of marking, in which DOM seems to be primarily determined by the need to differentiate the subject from the object. De Swart argues that in such languages, sentences with no semantic contrast between agent and patient must show a morphological contrast between them, in order to avoid ambiguity. The relevant semantic dimensions involve familiar prominence features, but are largely language-specific. In Malayalam (Dravidian), for example, object marking is mostly found on animate objects. However, in sentences which can in principle be interpreted in two different ways, it is also found on inanimate objects; the reason seems to be that without overt casemarking, the sentence would be ambiguous.

(7) a. kappal tiramaalakaḷ-e bheediccu
 ship.Nom waves-Acc split.Past
 ‘The ship broke through the waves.’

b. tiramaalakaḷ kappal-ine bheediccu
 waves.Nom ship-Acc split.Past
 ‘The waves split the ship.’

(de Swart 2007, citing Asher and Kumari 1997)

Such systems are called “global” because marking depends on properties of the subject, properties of the object, and the relation between them, along the lines of the discriminating/disambiguating view of marking. In contrast, “local” systems are those in which the presence of object marking is only dependent on the features of the object itself, along the lines of the coding/indexing approaches. As de Swart notes, global systems present a problem for Aissen’s model, since they depend on the simultaneous consideration of properties of the subject and object rather than properties of the object alone, but can be accounted for in de Swart’s analysis within the framework of Bidirectional Optimality Theory.

1.2.3 *DOM in transformational syntax*

Many analyses of phrasal syntax within the transformational paradigm assume two distinct positions for objects, VP-internal and VP-external, and postulate a correlation between object position and object marking (Diesing 1992, Dobrovie-Sorin 1994, van Geenhoven 1998, Torrego 1998, Ritter and Rosen 2001, Woolford 1999, 2000, 2001, de Hoop 1992, among others). The distinction is usually said to be semantically driven, dependent on a specific/referential vs. a nonspecific/nonreferential interpretation of the object. It is generally

assumed that VP-internal indefinite/nonspecific objects are syntactically less “visible” than VP-external definite/specific ones.

De Hoop (1992) observes that a morphological difference in case in languages with DOM is linked to a semantic difference in the interpretation of indefinites. Her analysis builds on work by Diesing (1992) on the interpretation of indefinite objects: Diesing suggests that arguments universally excluded from VP-internal positions bear the feature [+SPECIFIC], while VP-internal arguments are interpreted as nonspecific. This is formally modelled by requiring that NPs that are not overtly quantified must move out of VP before LF to escape existential closure. De Hoop (1992) assumes two categories of NPs, strong (presuppositional) and weak. Strong NPs are “anchors in conversation”; they are semantically characterised as generalised quantifiers and include referential, partitive, and generic expressions. Weak NPs are analysed as existentially quantified. Additionally, there are two kinds of object case: Weak Case, assigned VP-internally at D-structure, and Strong Case, assigned at S-structure to [Spec,AgrO]. A strong NP moves out of the VP to get Strong Case, while Weak Case requires syntactic adjacency to the verb. Overt accusative marking on VP-external objects is analysed as the morphological realisation of abstract Strong Case. This is exemplified, for instance, in Turkish, as shown in (1), where the marked and unmarked object receive specific and nonspecific interpretations, respectively.

These works deal only with casemarking, but since case and verbal agreement are treated as two aspects of the same phenomenon in this framework, roughly similar analyses have been proposed for differential object agreement. Both case assignment/checking (depending on one’s syntactic assumptions) and agreement are treated in terms of movement of the object to the specifier position of the relevant agreement head. In object-agreement languages, agreement serves as a specificity licenser, as argued by Sportiche (1995), among others.

Subsequent work has made it clear that specificity is not the only feature responsible for DOM. Adopting the premise that VP-internal and VP-external object positions may be associated with different semantic properties, Woolford (1999, 2000, 2001) aims to explain which objects occupy which of these positions and why, taking into account more complex patterns where there is no single semantic feature that triggers movement out of the VP and agreement. Instead, a VP-external object may have any of several distinct clusters of features.

The basic premise of Woolford’s proposal is that economy keeps objects in their base VP-internal positions unless that would violate certain Exclusion Principles. Exclusion Principles are modified versions of Diesing’s mapping principles, which exclude NPs bearing certain features from the VP-internal object position. Woolford assumes a family of Exclusion Principles based

on different semantic features including specificity, humanness, animacy, and number; on her view, these are separate principles and cannot be reduced to one more general principle. Each Exclusion Principle can be independently active in a language. The economy principle Avoid Movement, which prohibits moving objects out of their base position (Chomsky 1995), ensures that objects remain within the VP unless this would violate one or more Exclusion Principles. Cross-linguistic differences in restrictions on agreement are dealt with in terms of different rankings of various Exclusion Principles and the economy principle Avoid Movement.

In some languages DOM seems to depend on aspectual features of the verb. The idea that object marking correlates with aspect has been explored by Ramchand (1997), Meinunger (1998), Woolford (2000) and others, who suggest that the Asp head is involved in Case assignment to the object. Woolford (2000) claims that the Aspect node in Palauan supplies a Case feature when it is [+PERFECTIVE]. Ritter and Rosen (2001) provide a more sophisticated analysis, arguing that in languages with DOM the split in object marking is determined by the presence or absence of the feature [QUANT(ISATION)]. Quantised objects (their Class I objects) must check their QUANT feature, forcing such objects to move out of the VP and triggering agreement or casemarking, while non-agreeing or noncasemarked objects (their Class II objects) are not quantised and remain within the VP. The specific semantic contribution of the feature differs from language to language. When [QUANT] is an inherent feature of the verb, it has aspectual meaning: it encodes delimitedness or boundedness of the event. According to Ritter and Rosen (2001), this situation is exemplified in Finnish (Uralic), where objects of bounded events stand in the accusative case and objects of unbounded events take the partitive case; in Palauan, where object agreement correlates with boundedness of the event (as expressed through perfectivity); and in Mandarin Chinese, where the *ba*-construction is only possible with delimited events. On the other hand, when [QUANT] is a feature of the functional head Agr, it is uninterpretable and lacks inherent semantic content. In this case DOM is not sensitive to event type; instead, [QUANT] enters into a checking relation with definite/specific/animate objects, as in Turkish, Hebrew or Bantu. Ritter and Rosen (2001) do not explain why different semantic types are involved in object split in these languages, but emphasise the importance of treating [QUANT] as a feature which bears on the interpretation of verbs as well as objects, and which can be realised either as object case or object agreement.

1.3 Criteria for marking

As we saw in the previous section, most existing work on DOM appeals to inherent referential semantic features of the object noun phrase, such as ani-

macy, definiteness, or specificity, to distinguish marked and unmarked objects. Aissen's (2003a, 2003b) work is based on hierarchies of animacy and definiteness. For Næss (2004, 2006, 2007), the crucial notion of affectedness also correlates with referential properties of the object, although the correlation is indirect. Most transformational work on DOM is based on the premise that object marking patterns are defined in terms of semantic features (Diesing 1992, van Geenhoven 1998, and others). Work by Woolford (1999, 2000, 2001) relies on a family of Exclusion Principles based on specificity, humanness, animacy, and number. Ritter and Rosen (2001) use a general notion of boundedness which encompasses specificity and definiteness as well as event-boundedness.

These criteria are indeed useful in analysing patterns of DOM where objects that are characterised as semantically "strong" or "definite" show more agreement with the verb, or more casemarking, than objects without these properties. For example, semantic factors have been argued to provide the clearest explanation of patterns of DOM in Turkish or Hebrew, where the distribution of marked and unmarked objects is fairly straightforward and definable in terms of simple semantic features: Turkish marks specific objects, while Hebrew marks definite objects.³ Yet these factors do not adequately account for languages in which objects with the same semantic features can be either marked or unmarked. Such variation was noted by Bossong (1991:152), who concluded that in these languages the rules of DOM cannot be formulated precisely, but must allow for a certain degree of variability across speakers and situations.

Take, for example, Hindi. As discussed by Mohanan (1994:79 ff.), Singh (1994), and Aissen (2003b), the accusative marker *ko* in modern Hindi appears on specific human objects (in some dialects, all specific animate objects), whether definite or indefinite, but is impossible on indefinite inanimate nouns.

³However, von Heusinger and Kornfilt (2005) and Kornfilt (2009) argue that the correlation between specificity and accusative marking in Turkish is imperfect, and Danon (2006) claims that the definiteness condition on casemarking in Hebrew is a purely syntactic condition, related to the presence of an article, rather than a semantic property.

- (8) a. Ilaa-ne bacce-ko/*bacca uṭhaayaa
 Ila-Erg child-Acc/child lifted
 ‘Ila lifted the/a child.’
- b. Ravii-ne kaccaa kelaa kaaṭaa
 Ravi-Erg unripe banana cut
 ‘Ravi cut the/a unripe banana.’
- c. Ravii-ne kaccaa kele-ko kaaṭa
 Ravi-Erg unripe banana-Acc cut
 ‘Ravi cut the/*a unripe banana.’

(Aissen 2003b, from Mohanan 1994)

In other words, for specific human/animate nouns *ko*-marking is obligatory, while for inanimate indefinites, marking is disallowed. However, with inanimate/nonhuman definite objects, apparent optionality arises: such objects are either marked or unmarked.

On Aissen’s analysis of Hindi, the constraints that disallow *ko* on inanimate nonspecific objects and indefinite inanimates are dominated by *STRUC_C, which penalises a value for the morphological category CASE, so casemarking is absent. On the other hand, constraints that penalise the absence of case on human/animate objects dominate *STRUC_C, so accusative marking is required. Optionality of marking on inanimate/nonhuman definite objects is handled in terms of constraint reranking: the relative ranking of *STRUC_C against the constraints that penalise the absence of case on definite nonhuman/inanimate objects is undetermined. On this view, the patterns of grammatical marking are basically random and would be expected to reflect idiolectal and/or free variation.

On De Swart’s (2007) analysis, definiteness in Hindi does not trigger case-marking but is rather a **side effect** of the use of casemarking. De Swart claims that this eliminates optionality; however, his theory does not lead to straightforward and testable predictions for the distribution of *ko*, and it remains unclear how it can be formalised in an explicit grammar. Most importantly, no clear evidence is provided for the hypothesis that the presence of the accusative correlates with the speaker’s intention to “highlight” definiteness.

Many languages exhibit similar patterns of “optionality”. In Sinhala (Indo-Aryan), inanimate objects are never overtly marked, whereas animate objects are optionally casemarked.

- (9) mamə miniha(-wə) daekka
 I man-(Acc) saw
 ‘I saw the man.’ (Gair and Paolillo 1997:32)

In Kannada (Dravidian), accusative marking is obligatory on animate and/or definite objects but optional on indefinite inanimates (Lidz 2006). Casemarked indefinite objects are specific, while noncasemarked objects are ambiguous between a specific and nonspecific interpretation.

- (10) a. Naanu pustaka huDuk-utt-idd-eene
 I.Nom book look.for-Nonpast-be-1Sg
 ‘I am looking for a book/a certain book.’
 b. Naanu pustaka-vannu huDuk-utt-idd-eene
 I.Nom book-Acc look.for-Nonpast-be-1Sg
 ‘I am looking for a certain book.’ (Lidz 2006:11)

Lidz (2006) shows that the specific vs. nonspecific interpretation of noncase-marked objects correlates with their syntactic position, but not exactly in the way predicted by de Hoop’s (1992) analysis. Casemarked objects have a specific reading even when they occur inside the VP, while noncasemarked objects may occur VP-externally, in which case they must receive a specific reading.

Apparent optionality is also observed in languages with differential object agreement. In Ostyak (Uralic), object agreement may but need not appear with definite as well as indefinite objects, as shown in (11) (see Nikolaeva 1999, 2001 and the discussion of Ostyak in Chapter 8, Section 8.2):

- (11) tam kalaŋ we:l-əs / we:l-s-əlli
 this reindeer kill-Past.3SgSubj kill-Past-Obj.3SgSubj
 ‘He killed this reindeer.’
 a:muj kalaŋ we:l-əs / we:l-s-əlli
 some reindeer kill-Past.3SgSubj kill-Past-Obj.3SgSubj
 ‘He killed a/some reindeer.’

These examples demonstrate that in some languages, semantic features such as animacy, definiteness, or specificity alone are not enough to account for the distribution of agreement or casemarking on objects. Further, there is no evidence that the properties of the subject participant (as in “global” systems) or the degree of affectedness play a role. Optionality is observed with exactly the same subjects and exactly the same verbs.

1.4 Our proposal

We approach DOM from a different perspective. We believe that in many cases, the seeming unpredictability of variation in DOM is due to the limited set of features examined in previous work, and that a better account is possible if additional factors are taken into account. In particular, variations in marking can often be explained by reference to **information structure**, a level of sentence grammar where propositions, as conceptual states of affairs, are structured in accordance with the informational value of sentence elements and contextual factors. We propose a theory of DOM which emphasises the role of information structure in the marking patterns of objects. The information-structural distinction we make explains many of the unexplained areas in semantic analyses of DOM.

1.4.1 Information structure

Specifically, we propose that topicality plays an important role in DOM. Marked objects are often topical, while unmarked objects are nontopical. For example, in Ostyak (Uralic), object agreement is required for topical objects:

- (12) a. What did he do to this reindeer?
 b. tam kalaŋ we:l-s-əlli / *we:l-əs
 this reindeer kill-Past-Obj.3SgSubj kill-Past.3SgSubj
 ‘He killed this reindeer.’

Nontopical, focused objects disallow agreement:

- (13) a. Which reindeer did he kill?
 b. tam kalaŋ we:l-əs / *we:l-s-əlli
 this reindeer kill-Past.3SgSubj kill-Past-Obj.3SgSubj
 ‘He killed this reindeer.’

This is true independent of semantic features of animacy and definiteness; in both of these examples, the object is animate and definite.

Following Lambrecht (1994) and others, we understand topicality as a pragmatic relation that holds between a referent and the proposition expressed by an utterance: topicality has to do with the construal of the referent as pragmatically salient, so that the assertion is made **about** this referent. Topicality is not an inherent property of a referent, and although it correlates with the role played by the referent in the preceding discourse, the correlation is imperfect. It cannot be unambiguously established on the basis of the referential features of the object either; rather, it depends on the speaker’s assessment of

its saliency within a given **communicative context**, as we discuss in Chapter 3. Since we view information structure as part of grammar, we treat grammaticality as including pragmatic well-formedness. For instance, object agreement is ungrammatical in (13) only if it construed as an answer to the question *Which reindeer did he kill?*, though it is possible in the context in (12). In order to demonstrate the pragmatic role of a referent, we provide examples in context whenever possible.

In some respects, our view most closely resembles the proposal of Næss (2004): like ours, her analysis relies not on the semantic features of the object, but rather on its perceptual role. In fact, Næss (2004) suggests that affectedness, which plays a crucial role in her treatment of DOM, has two aspects: expressing a part/whole relation, and saliency. Saliency is a necessary property of topics, so the two notions, though distinct, seem to be closely related (see Croft 1991:155 for further discussion of the correlation between affectedness and topicality). However, for Næss (2004) saliency has to do with which entities are of more interest for human perception in general. Her analysis does not, then, address contextual factors. On our approach, saliency relates to the role played by a referent in the pragmatic structure of the proposition, established within a given communicative context.

Our proposal provides a different functional motivation for DOM from the discriminatory/disambiguating perspective: we argue that DOM was originally motivated by the need to highlight **similarities** between subjects and topical objects, which tend to be grammatically marked, as opposed to nontopical objects. That is, it arises from the need to give overt expression to a property that is **common** to subjects and (some) objects but less typical of other non-core grammatical functions: languages tend to mark topics, whether subjects or objects, either by agreement or by casemarking. This means that topical objects are not functionally marked, atypical, or uncommon; in fact, we suggest that objects are just as likely to be topical (grammatically marked) as nontopical (focused, unmarked).

The idea that topicality may play a role in DOM has been mentioned in previous work (Aissen 2003b, Leonetti 2003, and others). However, it has not received extensive elaboration in the family of functionally-typological and Optimality Theoretic analyses of DOM, nor has it been standardly incorporated into most existing generative or transformational treatments. Some research employs the notion of **prominence** interpreted more broadly than Aissen's Prominence Scales (de Hoop and Narasimhan 2005, de Hoop and Lamers 2006, de Hoop 2008). De Swart (2007:138) defines prominence as a feature that "is concerned with the centrality of an entity in the discourse or with the readiness with which an entity presents itself to the speaker as a topic of conversation". His analysis distinguishes two levels of prominence: (i) discourse prominence, which reflects the status of the argument in the discourse, and (ii)

semantic prominence, which depends on the intrinsic semantic features of the argument. The notion of discourse prominence seems to be comparable with our notion of topicality; however, de Swart does not provide a detailed characterisation of discourse prominence, noting only that semantic prominence influences discourse prominence because an inherently prominent element is more likely to be topical.

Another “exception” is Meinunger (1998): working within the Minimalist Program, he suggests that the distinction between topical and nontopical objects is interpreted syntactically, and that the semantic feature that all VP-external phrases have in common is [+TOPIC]. This feature activates VP-external agreement projections, so that topical NPs must move to a SpecAgr position for Case checking. He also assumes an inherent relationship between verbal agreement and Case, which are seen as two aspects of the same phenomenon, in that both are effects of the movement of an argument to the specifier position of an agreement head. Languages differ in how they mark scrambled topical objects: some languages mark topical and nontopical objects with morphologically different cases, while in other languages the difference in the interpretation of the objects is linked to object agreement, and in a third type of language the difference is only positional. Meinunger views topical arguments as discourse-linked/familiar and contextually salient. In addition, the speaker must intend to construe them as topics, as opposed to presentational focus or comment, which normally contains discourse-new elements. However, his definition of topic and the relationship between topicality and the referential features of the object are not clearly spelt out. For instance, Meinunger claims that topicality is realised as specificity in Turkish and animacy in Hindi, but does not explain this difference, though object marking clearly depends on very different semantico-pragmatic conditions in these languages.

Escandell-Vidal (2009) presents a detailed analysis of DOM by means of the preposition *a* in Balearic Catalan. Although differential object casemarking in many varieties of Catalan is explained by appeal to semantic features of animacy or definiteness, in Balearic Catalan referential properties of the object play a secondary role. Pronominal objects are always casemarked, but for lexical objects casemarking is (partly) determined by topicality. As in our analysis, Escandell-Vidal understands topicality as “aboutness”, and therefore Balearic Catalan seems to behave very similarly to a number of the languages we discuss in this book. Her analysis differs from ours in that she does not discuss the alignment of grammatical functions of marked and unmarked objects with information structure roles, though she shows that there are positional differences between topical and nontopical objects, since Balearic Catalan is a language with designated topic and focus positions. Topical objects cannot occur in the canonical focus position, and instead must be either left- or right-dislocated. However, only discourse-old topics, and not all topical objects,

are casemarked. This implies that both marked and unmarked objects can occur in topic position, and raises the question of whether they display identical syntactic behaviour.

1.4.2 Syntax

This leads us to the next point. Our view differs from most previous proposals in another respect: we believe that existing typological analyses of DOM, which mainly concentrate on its functional motivation and the cross-linguistic distribution of morphological marking, do not pay sufficient attention to the syntactic side of the phenomenon. The behavioural properties and syntactic status of marked vs. unmarked objects are rarely discussed, and to our knowledge, there are few systematic cross-linguistic studies of the differences between them.

We have found that in some languages marked and unmarked objects do not differ syntactically and, arguably, realise the same grammatical function, the object. In other languages they not only have different information structure roles, but also exhibit different behavioural syntactic profiles. For example, marked and unmarked objects in Ostyak differ in their ability to control coreference with the subject of an action nominal dependent clause, allow possessor topicalisation, control possessor reflexivisation, and launch floated quantifiers. In other languages as well, a number of syntactic tests distinguish marked objects and unmarked objects, with the marked object displaying a larger number of properties associated with core arguments. We suggest that in these languages, marked and unmarked objects bear different **grammatical functions**. Our proposal is cast within the framework of Lexical Functional Grammar (LFG), which distinguishes between the primary object **OBJ** and the secondary or semantically restricted object **OBJ_θ** (Bresnan and Kanerva 1989). These two types of objects are usually discussed in connection with double object constructions, but in many languages, both types are available in single object constructions as well, with the choice between them determined by their information structure role.

As mentioned above, many analyses of DOM assume two distinct positions for objects, VP-internal and VP-external, and posit a correlation between the position of the object and grammatical marking (Diesing 1992, Dobrovie-Sorin 1994, van Geenhoven 1998, Torrego 1998, Ritter and Rosen 2001, Woolford 1999, 2000, 2001, de Hoop and Malchukov 2007, among others). Issues of word order and the positions of objects have been studied thoroughly within the LFG framework by Butt and King (1996), Choi (1999) and many other researchers: their work has clearly shown that information structure role can be relevant for word order constraints, and that information structure has a strong effect on where objects can appear. However, we do not posit a di-

rect relation between DOM and the phrase structural position of the object in all languages, because at least in some of the languages we analyse there is no obvious connection between the position of the object and its grammatical marking. More generally, we do not assume that grammatical functions or information structure roles must be identified configurationally. LFG analyses grammatical functions as syntactic primitives, and does not define them in terms of phrase structure position. Positional generalisations concerning the behaviour of different grammatical functions or arguments bearing different information-structure roles can be easily modelled within LFG's **projection architecture** (Kaplan 1987, Asudeh 2006), which allows constraints to refer to and relate different aspects of the structure of an utterance.

1.4.3 Diachrony and grammaticalisation

Our analysis also provides a diachronic explanation for why referential semantic features such as animacy, definiteness, and specificity often play a role in DOM. We propose that these features are involved in grammaticalisation of topichood.

Abundant diachronic evidence shows that DOM often originates as a marking device for topics. We take this situation to be historically primary. One possible direction of change involves widening of topical marking, where marking spreads to certain nontopical objects. In a language with topical object marking, the marking can generalise or extend, applying to nontopical objects which have features typical of topics. When this happens, marking patterns become automatic consequences of distinctions at other levels of structure. As a result, the role of information structure in object marking is diminished: marking is obligatory not only for topical objects, but also for objects with certain semantic features, independent of their information structure role.

The opposite direction of change involves narrowing of topical marking: marking becomes specialised for topics which bear certain semantic features. In the relevant languages, only a subset of topical objects are formally marked, while nontopical objects remain unmarked. Narrowing usually involves the most typical members of the set. Objects ranked high on the prominence hierarchies are frequent topics, so topical marking can become restricted to them. At the next stage of grammaticalisation, the connection to information structure may be completely lost, so that object marking becomes dependent on semantic features alone.

Thus, we propose that different patterns of DOM arise as a result of different directions of grammaticalisation of topic marking on objects. Features that are typical of topics come to be required, or sufficient, for object marking.

1.4.4 Limits of our analysis

Our theory of DOM does not extend to all instances of variable marking of objects. Since we are interested in differences in interpretation that depend on the **presence or absence** of grammatical marking, we only consider languages where a subset of objects is grammatically marked (by either case or agreement), while another subset receives no marking at all, as in the examples cited above. We will not address those instances of DOM where objects can take two alternative casemarkers, as attested, for example, in Icelandic (North Germanic):

- (14) a. Hann klóraði mig
 he.Nom scratched me.Acc
 ‘He scratched me.’
- b. Hann klóraði mér
 he.Nom scratched me.Dat
 ‘He scratched me.’ (Næss 2004, citing Barðdal 2001)

The Icelandic accusative-dative alternation has been reported to reflect a semantic contrast: in (14a) the scratching is perceived as a forceful act of violence, possibly painful for the patient participant, while in (14b) the interpretation involves volitionality on the part of the patient. In other languages, such as Russian (Slavic) or Finnish, the case of the object differs depending on aspectual characteristics of the verb. As interesting as such instances may be, we believe that they deserve separate treatment and do not fall under the same generalisations as DOM in languages which contrast formally marked and unmarked objects.

The focus of this book is the grammatical function of **object**. We limit the scope of our study to object marking — casemarking or agreement — in order to more fully explore the relation between the presence and absence of marking and the status of object arguments at different levels of linguistic representation. We exclude languages where some patient/themes are not syntactic objects at all, but participate in various detransitivising constructions, and may remain syntactically unexpressed or incorporated into the verb. In other words, we only deal with proper syntactic arguments. In this way, our work differs from some previous research on DOM such as, for example, Næss (2004), which treats alternative encodings of **patient/theme arguments** independently of their syntactic status.

Furthermore, our analysis does not address the distribution of verbal clitics or incorporated pronouns, although in many languages they are associated with some but not all objects, as discussed by Bresnan and Mchombo (1987)

for Chicheŵa (Central Bantu), Culy (2000) for Takelma (Penutian), Bowers (2004) for Bardi (Nyulnyulan), Donohue (2004) for *Tukang Besi*, and Jaeger and Gerassimova (2002) and Jaeger (2004) for Bulgarian (Slavic), among others. The status of object markers in a number of Bantu languages is debatable (see, for example, Seidl and Dimitriadis 1997 and Morimoto 2002 on Swahili, and Woolford 2001 on Ruwund), so we will not discuss such languages unless we can demonstrate that the marking on the verb is actually agreement marking and not pronominal incorporation (see Bresnan and Mchombo 1987 and Chapter 2, Section 2.6 for detailed discussion of this issue).

We also limit our scope to the examination of nominative-accusative languages, and have little to say about languages with an ergative-absolutive case-marking system, where the object is absolutive. Malchukov (2006) notes that DOM is more typical of nominative-accusative systems, and proposes a principle of argument marking that holds of both types of languages: languages tend to avoid manipulation of the casemarking of the unmarked argument, i.e. the nominative argument in nominative-accusative systems, and the absolutive argument in ergative-absolutive systems. Deemphasis of the absolutive object tends to give rise to voice alternations such as antipassive, rather than case alternations. We do not take a stand here on whether this view is correct, but the fact remains that DOM is more frequent in nominative-accusative languages, and we will concentrate on such languages in this book.

1.5 Structure of the book

The book is structured as follows.

Chapter 2 provides a brief introduction to Lexical Functional Grammar (LFG) (Bresnan 2001, Dalrymple 2001, Falk 2001), the syntactic framework employed in this book. Chapter 3 presents an overview of the model of information structure which we assume, and in Chapter 4 we present our proposal for the treatment of information structure in LFG.

In Chapter 5 we discuss the expression of primary topics in syntax, suggesting that although there is no unique alignment between information structure roles and grammatical functions, there are important cross-linguistic tendencies in the grammatical expression of primary topics.

Chapter 6 addresses languages in which grammatical marking for nonsubject topics occurs with a variety of syntactic roles: it is possible for objects, possessors, instruments, and other syntactic functions. Comrie (2003) discusses agreement patterns in such languages under the rubric of “trigger-happy agreement”, and we show that casemarking can work in a similar manner. These data constitute the main evidence for the relevance of topichood in the grammatical marking of nonsubjects.

Differential object agreement and differential object casemarking are dealt with in Chapter 7, where we examine languages in which grammatical marking of nonsubject topics is syntactically restricted to objects. In such languages, (a subset of) topical objects receive special agreement or casemarking, while (a subset of) nontopical/focused objects remain unmarked.

We then take a closer look at syntactic differences between marked and unmarked objects in languages in which grammatical marking is confined to objects, and discuss cross-linguistically typical patterns of alignment between syntax and information structure. We show in Chapter 8 that in some such languages, a number of syntactic tests distinguish marked objects and unmarked objects, with the marked (topical) object displaying more properties associated with core grammatical functions. Chapter 9 suggests some generalisations concerning ditransitive constructions in languages with DOM.

Chapter 10 proposes a historical explanation for semantic patterns of DOM, which can arise as a result of different directions of grammaticalisation of nonsubject topic marking. Since features such as definiteness, specificity or animacy are known to be associated with topics, grammatical marking may spread to nontopical objects with these semantic features, or may narrow to include only topical arguments which also have these features.

Chapter 11 concludes the book.

Syntactic assumptions

Nontransformational, constraint-based theories of grammar such as Construction Grammar (Croft 2001, Goldberg 2006), HPSG (Sag et al. 2003), Role and Reference Grammar (Van Valin 2003), Simpler Syntax (Culicover and Jackendoff 2005), and Lexical Functional Grammar (Bresnan 2001, Dalrymple 2001, Falk 2001) represent different aspects of the structure of an utterance as separate but related grammatical modules. Such theories assume that syntactic structure is related to semantics, information structure, and other linguistic levels not by means of transformational operations, but by constraints involving one or more levels of structure. Information structure roles may be associated with particular phrase structural positions, but these positions do not define the roles (as they often do in tree-based, transformational theories such as Principles and Parameters Theory or Minimalism: Chomsky and Lasnik 1993, Chomsky 1995). Instead, the relations between grammatical, semantic, and information structural roles are specified in terms of constraints involving different levels of linguistic representation.

In LFG, these different aspects of linguistic structure — phrase structure, grammatical functions, information structure — are represented by structures that may be of different formal types. The phrasal structure of the sentence is represented by a phrase structure tree, the **constituent structure** or **c-structure**. Grammatical functions like subject and object are represented by the **functional structure** or **f-structure**. **Information structure** (Choi 1999, Butt and King 1996, 2000) is related to other grammatical levels within the **projection architecture** of LFG (Kaplan 1987, Asudeh 2006). LFG provides the tools needed to analyse the relation between grammatical functions and information structure roles and to provide a formal treatment of the grammar of object marking and its relation to semantics and information structure.

2.1 Grammatical functions

A central tenet of LFG theory is the dual nature of syntactic structure. The abstract grammar of all languages is organised around syntactic relations like subject, object, and adjunct: all languages distinguish arguments from adjuncts, and subjects from nonsubjects, for example. These abstract grammatical relations are overtly expressed in phrasal structure in very different ways in different languages. Some languages have rigidly fixed word order, while others allow a great deal of freedom in word order, and languages vary in terms of where the verb must appear in relation to its arguments. The level at which word order and phrasal constituency is represented has its own internal organisational principles, different from the level at which abstract grammatical relations are represented. The two syntactic structures of LFG, constituent structure and functional structure, reflect this duality.

2.1.1 Grammatical function diagnostics

LFG assumes that abstract grammatical functions such as subject, object, complement, and adjunct are theoretical primitives, not defined in terms of other levels of structure. That is, the subject of a sentence is not defined as a phrase appearing in a particular phrase structural position or as an entity bearing a particular semantic role; instead, subject, object, and other grammatical functions are basic concepts of the grammar, around which various grammatical properties tend to cluster (Falk 2006).

In LFG, diagnostics identifying particular grammatical functions appeal to abstract behavioural grammatical properties defined at functional structure. In many languages, for example, the subject is the argument which agrees with the verb, or is the antecedent of a reflexive, or is absent in a subordinate clause under identity with a matrix clause argument. Mohanan (1994) shows that binding relations for pronouns in Hindi are constrained by the grammatical function of the antecedent. The antecedent of a pronoun cannot be the subject of its clause:

- (1) Vijay ne Ravii ko uskii saikil par biṭhaayaa
 Vijay Erg Ravi Obj his bicycle Loc sit.Caus.Perf
 ‘Vijay_i seated Ravi_j on his_{*i,j} bike. (Mohanan 1994:126)

Matsumoto (1996) discusses honorification in Japanese, showing that some verbs take special forms to honor the subject referent, and not other grammatical functions. The verb form *o-V ni naru* is used to honor the subject *sensei* ‘teacher’, and cannot be used to honor a nonsubject:

- (2) a. sensei wa hon o o-yomi ni narimashita
 teacher Top book Acc Honorific-read Cop become-Polite-Past
 ‘The teacher read a book.’ (Matsumoto 1996:27)
- b. *Jon wa sensei no o-tasukerare ni natta
 John Top teacher by Honorific-help Cop become-Past
 ‘John was saved by the teacher.’ (Matsumoto 1996:28)

None of these properties is definitional of subjecthood, and indeed all of these properties have been found to be associated with nonsubjects as well as subjects, depending on the language. In fact, these properties tend to target the top of the grammatical function hierarchy in (3), originally proposed by Keenan and Comrie (1977), and so distinguish subjects from other arguments in some languages, and core arguments from noncore arguments, or arguments from adjuncts, in other languages.

- (3) subject > object > oblique > adjunct

Because the subject is the argument that stands at the head of the grammatical function hierarchy, it is often the only grammatical function that can participate in certain processes, and can be reliably identified in this way.

Diagnostics targeting nonsubject grammatical functions, specifically objects, also vary from language to language. For example, we have seen that some languages have object agreement, which can function as a test for objecthood, and some languages have a special relativisation strategy for objects that is different from the strategy used for other grammatical functions.

Diagnostics for grammatical functions that are stated in terms of nonsyntactic criteria are more difficult to use, and less reliable. Unlike transformational theories, LFG does not assume that grammatical functions are defined in terms of positions in a phrase structure tree. Thus, in many languages the appearance of a phrase in a particular position is not criterial for identification of the phrase as bearing a particular grammatical function, although in some fixed word order languages, a particular phrase structure position may be reserved for a particular grammatical function. Grammatical function alternations are also unreliable tests; it is often claimed that objects can be identified as those arguments of an active verb which appear as the subject of the corresponding passive verb, but although this is often true, there are passive subjects which are not objects of the active verb (*Bill is said to be available*!/**Someone said Bill to be available*). For more discussion of grammatical function diagnostics, see Andrews (1985) and, for objects, Hudson (1992); for an LFG perspective, see Dalrymple (2001: Chapter 2) and, for objects, Börjars and Vincent (2008) and Lam (2008).

2.1.2 *Grammatical functions and semantic roles*

Predicates encode an association between grammatical functions and particular semantic roles: for example, the active version of the verb *kick* associates its subject with the semantic role of kicker, and its object with the semantic role of the entity that is kicked (*The man kicked the ball*). **Mapping theory** (Levin 1986, Bresnan and Kanerva 1989, Butt 1995, Alsina 1996) is a theory of the relation between **argument structure** (the level at which semantic roles are represented) and functional structure (the grammatical functions associated with those roles).

Researchers in mapping theory agree that argument structure encodes a certain amount of semantic information about the arguments of a predicate, but different researchers have adopted different views of the content and representation of argument structure. Much research is based on the assumption that argument structure makes available a set of abstract semantic roles like agent, source, theme, patient, and goal, which are ranked relative to one another on a **thematic hierarchy**. Bresnan and Kanerva (1989) propose the thematic hierarchy shown in (4):

- (4) Thematic hierarchy (Bresnan and Kanerva 1989):
- AGENT > BENEFACTIVE > RECIPIENT/EXPERIENCER
> INSTRUMENT > THEME/PATIENT > LOCATIVE

These semantic roles are systematically related to grammatical functions (see, for example, Bresnan and Kanerva 1989 and Bresnan and Zaenen 1990): on this view, mapping theory relates the agent role (kicker) of *kick* to SUBJ, and the patient role (object kicked) to OBJ. Other versions of mapping theory adopt a different view of argument structure, often in an attempt to define or provide a clearer characterisation of semantic roles. Butt (1995) assumes Jackendoff's (1983, 1990) Conceptual Semantics representation of argument structure, which involves the use of a set of conceptual primitives such as CAUSE, GO, and BE to encode basic semantic relations. Alsina (1996) works within a version of the proto-role theory of Dowty (1991), which provides a set of entailments classifying the arguments of a predicate according to the degree to which they exhibit proto-agent or proto-patient properties. Our theory of argument marking does not depend on the details of these theories of argument structure, and so we will appeal when necessary to semantic roles like agent, patient, theme, source, and goal, without presupposing any particular theory of their representation or definition at argument structure.

Whatever theory of argument structure is adopted, all LFG researchers agree that mapping theory must include several very general constraints on the relation between argument structure and functional structure. The first is **function-argument biuniqueness** (Bresnan 1980), which requires each semantic role to

correspond to exactly one grammatical function, and no grammatical function to correspond to more than one semantic role. For example, there are no predicates whose argument structure contains two different semantic roles (an agent and a theme, for example) with both roles mapped to a single grammatical function, and there are no predicates whose argument structure contains a single semantic role mapped to two different grammatical functions:

(5) Function-argument biuniqueness:

Argument structure:	ALLOWED: θ_1 θ_2	DISALLOWED: θ_1 θ_2	DISALLOWED: θ_1
Functional structure:	$\begin{array}{c} \\ \text{GF}_1 \end{array}$ $\begin{array}{c} \\ \text{GF}_2 \end{array}$	$\begin{array}{c} \diagdown \quad \diagup \\ \text{GF}_1 \end{array}$	$\begin{array}{c} \diagup \quad \diagdown \\ \text{GF}_1 \quad \text{GF}_2 \end{array}$

Another general constraint is the **subject condition**, which stipulates that a verbal predicate must have a subject. This is a more controversial condition which has been challenged by some researchers (for discussion, see Berman 1999), and it may be that the subject condition holds in only some languages rather than being a universal principle of mapping theory. Though (as far as we know) the subject condition holds in all of the languages that we will discuss, we will not take a position on the universality of the subject condition in this work.

2.2 Objects

Since our focus is the differential marking of objects, we will be most concerned with object grammatical functions. Here, we provide a brief overview of the standard LFG view of objects. Recent LFG work has explored the object functions from a variety of perspectives: see Lam (2008) for a thorough discussion of object functions with a focus on constructions with multiple objects, and Börjars and Vincent (2008) for an LFG-based perspective that differs from the traditional view in interesting ways, particularly with regard to the analysis of the English double object construction. We return to a discussion of the object functions and object mapping in Chapter 9.

LFG assumes that more than one object function can be manifested in a language. In early LFG, these object functions were called OBJ (object) and OBJ2 (second object) (Bresnan 1980, Kaplan and Bresnan 1982). In a double object construction, the object which shares more properties with the object of a monotransitive verb was called the (primary) object, OBJ, and the other argument was called the second object, OBJ2.¹

¹Dryer (1986) uses the terms **primary object** and **secondary object** to draw nearly the same distinction; see Chapter 9 for further discussion of LFG's object functions in double object constructions.

Example (7a) shows the monotransitive verb ‘eat’ with a single object, the class 7 noun ‘food’. With the addition of the applicative affix, a benefactive/malefactive applied OBJ ‘wife’ is added, and ‘food’ becomes a secondary OBJ_{THEME}.

Bresnan and Moshi (1990) show that in Kichaga, applicative constructions with more than two objects are possible, establishing conclusively that a simple two-way distinction between OBJ and OBJ₂ is not sufficient for grammatical description. Consider example (8):

- (8) n-ǎ-l'é-kú-shí-kí-kóŕ-í-à
 Foc-1Subj-Past-17Obj-8Obj-7Obj-cook-Applicative-FinalVowel
 ‘She/he cooked it with them there.’ (Bresnan and Moshi 1990:151)

This example contains three object markers, representing a locative object, an instrumental object, and a patient object. The grammatical structure of this example cannot be analysed using only the grammatical functions OBJ and OBJ₂, since this example contains three object-like functions. Bresnan and Moshi show that in (8), the instrumental OBJ is the unrestricted OBJ, while the locative and patient arguments bear semantically restricted OBJ functions OBJ_{LOC} and OBJ_{PATIENT}.

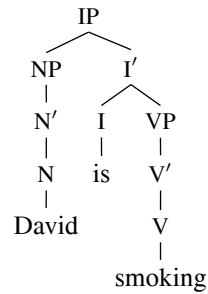
In much previous LFG literature, the OBJ/OBJ_θ distinction was assumed to play a role primarily or exclusively in the analysis of ditransitive constructions, but we will see in Chapter 8 that the distinction is also relevant for monotransitive predicates. For discussion of other languages in which monotransitive verbs can take either OBJ or OBJ_θ, see Çetinoğlu and Butt (2008) for Turkish and Dahlstrom (2009) for Meskwaki (Algonquian).

2.3 Levels of syntactic representation

The following brief sketch presents the formal syntactic concepts that will be relevant in our analysis of DOM; for a more complete introduction to LFG, see Bresnan (2001), Dalrymple (2001), and Falk (2001).

The constituent structure tree or **c-structure** represents linear order and phrasal grouping:

(9) David is smoking.



This tree represents the word order and phrasal constituency of the words that actually appear in the sentence. This tree is not the product of a movement process or the input to any transformational operation; LFG does not assume the existence of processes that modify or destroy grammatical structure.

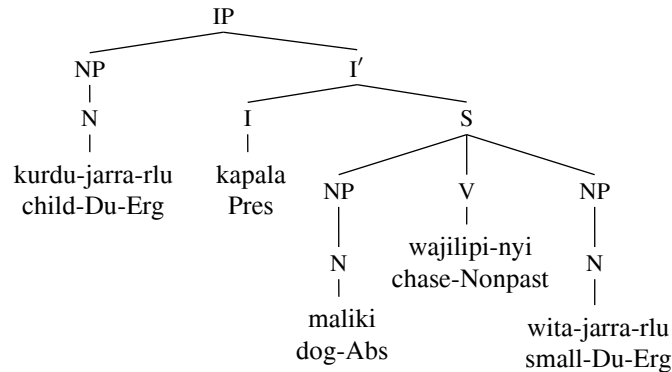
The categories in the tree shown in (9) are standard X' -theoretic, endocentric categories. We have represented the name *David* as a NP, but since the internal structure of nominal phrases is not of particular concern here, we will not distinguish between NP and DP.³ All phrase structure categories in English are endocentric (that is, headed). Some languages also make use of the exocentric category S in their phrase structure (Kroeger 1993, Bresnan 2001). According to Bresnan (2001:110), S is a non-headed category which is not subject to normal X' -theoretic constraints on c-structure configurations or the relation between c-structure and f-structure. Not all languages use this category, and many of those that do have relatively free word order. The Australian language Warlpiri has notoriously free word order (Simpson 1991); Austin and Bresnan (1996) give the following c-structure for a typical Warlpiri sentence:

³See Börjars et al. (1999) for discussion of functional categories and the NP/DP distinction in LFG.

- (10) kurdu-jarra-rlu kapala maliki wajilipi-nyi wita-jarra-rlu
 child-Du-Erg Pres dog.Abs chase-Nonpast small-Du-Erg

‘The two small children are chasing the dog.’

(Austin and Bresnan 1996:225)



Unlike English, the specifier position of IP in Warlpiri is not associated with the subject function, and any constituent can appear there. Since the daughters of S can be of any category and can be reordered freely, all orders of words in this sentence are acceptable, so long as the auxiliary appears in second position (Simpson 1991, Austin and Bresnan 1996).

The functional structure or **f-structure** represents syntactic predicate-argument structure as well as features like case, person, number, gender, tense, and aspect, as shown in (11). This is the level at which grammatical functions like subject and object are represented.

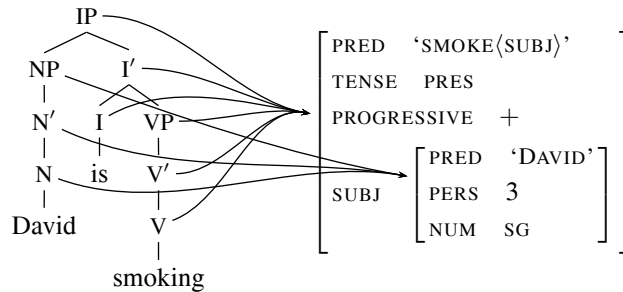
The f-structure for the sentence *David is smoking* is given in (11). The f-structure for the subject *David* appears as the value of the feature SUBJ, and contains three features: PRED, whose value is the main syntactic predicate ‘DAVID’ of the subject phrase; the person feature PERS with value 3; and the number feature NUM with value SG, or singular. The f-structure for the entire sentence has the feature PRED with value ‘SMOKE<SUBJ>’, indicating that the syntactic predicate SMOKE requires one argument, the SUBJ(ect). It also has a TENSE feature with value PRES(ent), and a PROGRESSIVE feature with value +. LFG researchers generally follow a convention of representing only the f-structure information that is relevant for the point at hand, and so the f-structures we present will often contain only information that is necessary for purposes of the current discussion, and will omit all other features.

(11) David is smoking.

$$\left[\begin{array}{l} \text{PRED 'SMOKE(SUBJ)'} \\ \text{TENSE PRES} \\ \text{PROGRESSIVE +} \\ \text{SUBJ } \left[\begin{array}{l} \text{PRED 'DAVID'} \\ \text{PERS 3} \\ \text{NUM SG} \end{array} \right] \end{array} \right]$$

LFG assumes a piecewise relation between levels of linguistic structure, so that parts of one level of structure are related to the corresponding parts of another level. For example, the parts of the c-structure tree that make up the subject phrase are related to the SUBJ f-structure by a correspondence function called ϕ . The constituent structure tree and functional structure for the sentence *David is smoking* are given in (12), with the ϕ relation between nodes of the c-structure tree and f-structures indicated by arrows.

(12) David is smoking.



Formally, the arrows from c-structure to f-structure represent a function relating nodes of the c-structure to parts of the f-structure. In (12), the IP, I', I, VP, V', and V nodes of the tree are all related to the f-structure for the entire sentence, and the NP, N', and N nodes are related to the f-structure for the subject. This configuration adheres to the general principles governing the relation between c-structure and f-structure presented by Bresnan (2001:102):

- (13) a. C-structure heads are f-structure heads. [In (12), the IP, I', and I are mapped to the same f-structure; the same is true for the VP, V', and V, as well as for the NP, N', and N.]

- b. Specifiers of functional categories are the grammaticalised discourse functions $_{DF}$ (including $_{SUBJ}$). [In (12), the NP in the specifier of IP is mapped to the grammatical function $_{SUBJ}$.]
- c. Complements of functional categories are f-structure co-heads. [In (12), the I and the VP are mapped to the same f-structure.]

Other linguistic levels are related to c-structure and f-structure in a similar way, by means of piecewise functions from parts of one structure to parts of another. These functions are called projection functions, and because of this, the different levels are also called projections (Kaplan 1987, Asudeh 2006). In Chapter 4, we will see how information structure is represented, and how it is related to other linguistic levels by means of a projection function. For more discussion of the c-structure/f-structure mapping and how it is constrained, see Bresnan (2001), Dalrymple (2001), and Falk (2001).

2.4 Describing linguistic structures

In LFG, the grammar of a language is a complex description of possible grammatical structures and the relations between them. The phrase structure rules of a language determine the acceptable constituent structure trees of the language (as originally proposed by McCawley 1968); annotations on those trees define the f-structures that correspond to them and, as we will see in Chapter 4, other levels as well, including information structure. In the following, we confine attention to c-structure and f-structure, and show how to define and constrain the relation between them.

Consider the following portion of the c-structure and f-structure given in (12):



The rule in (15) licenses the c-structure configuration in (14), requiring it to consist of a mother IP node and daughters NP and I':

$$(15) \text{ IP} \longrightarrow \text{NP I}'$$

We also require a specification of the relation between this phrasal configuration and the f-structure it corresponds to. In particular, IP and I' must correspond to the same f-structure (that is, the I' is the **f-structure head** of the IP), and the NP corresponds to the $_{SUBJ}$ of that f-structure. We can impose this requirement with the following annotated rule:

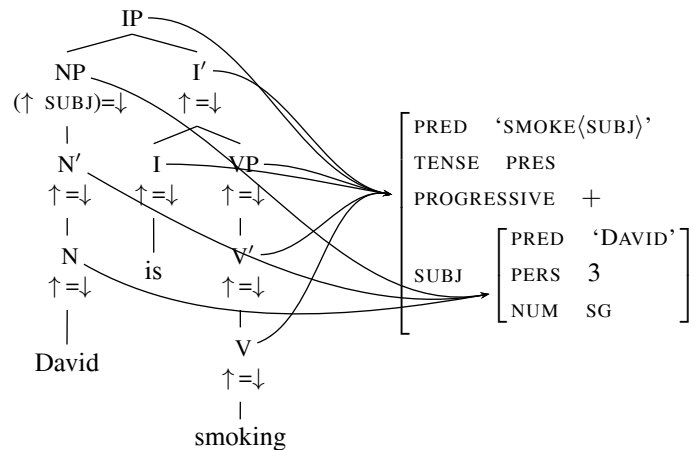
$$(16) \text{ IP} \longrightarrow \begin{array}{cc} \text{NP} & \text{I}' \\ (\uparrow \text{ SUBJ})=\downarrow & \uparrow=\downarrow \end{array}$$

In an annotated phrase structure rule, the up arrow \uparrow refers to the f-structure corresponding to the mother node, and the down arrow \downarrow refers to the f-structure of the daughter node that it appears on. Thus, the annotation $(\uparrow \text{ SUBJ})=\downarrow$ on the NP means that the f-structure corresponding to the mother IP, \uparrow , has a SUBJ feature whose value is the f-structure corresponding to the daughter NP, \downarrow . In other words, the NP that appears in the specifier of IP corresponds to the subject of the IP. This rule adheres to principle (13b) governing the relation between c-structure and f-structure, and is correct for English, a configurational language which associates the specifier of IP position with the subject function. In other languages, the specifier of IP may be associated with other functional roles; see Bresnan (2001) and Dalrymple (2001) for more discussion.

The annotation $\uparrow=\downarrow$ on the I' means that the f-structure corresponding to the IP (\uparrow) is the same as the f-structure corresponding to the I' (\downarrow). In other words, the I' is the f-structure head of the IP and shares all of its f-structure information. This follows principle (13a) governing the relation between c-structure and f-structure, which requires c-structure phrases and their heads to correspond to the same f-structure.

The remaining nodes in the phrase structure tree are annotated with $\uparrow=\downarrow$, requiring them to have the same f-structure as their mother. This follows from principle (13c), which requires the VP and the I' to correspond to the same f-structure, and from principle (13a), which requires a phrase and its head daughter to correspond to the same f-structure.

(17) David is smoking.



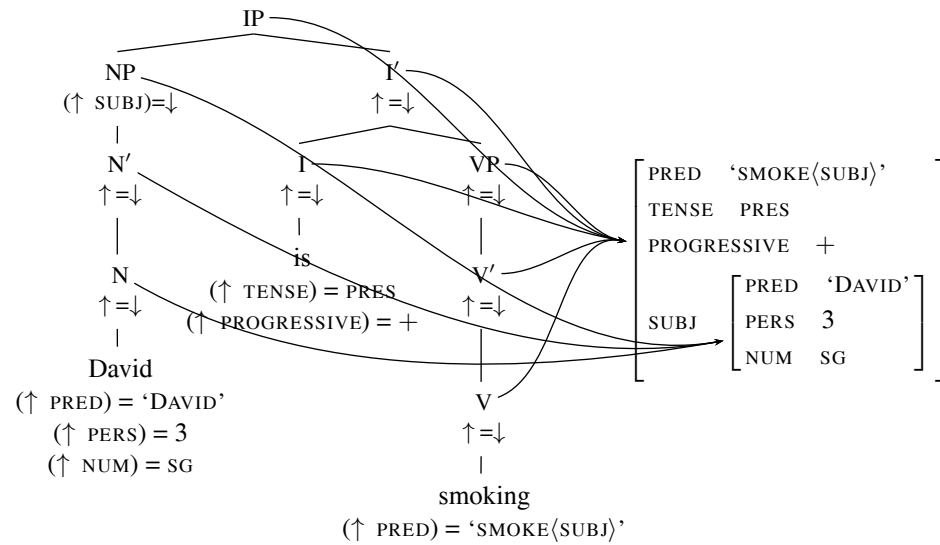
Lexical entries are annotated with information about the preterminal node that dominates them: *David* specifies information about the preterminal N, *is* specifies information about the preterminal I, and *smoking* specifies information about the preterminal V. Preliminary lexical entries for *David*, *is*, and *smoking* are given in (18):⁴

(18) Preliminary lexical entries for *David*, *is*, and *smoking*:

- David* (↑ PRED) = 'DAVID'
(↑ PERS) = 3
(↑ NUM) = SG
- is* (↑ TENSE) = PRES
(↑ PROGRESSIVE) = +
- smoking* (↑ PRED) = 'SMOKE⟨SUBJ⟩'

Adding this information to the tree gives us the following configuration, which shows where the information in the f-structure is introduced and how it flows through the tree:

(19) David is smoking.



⁴For expository purposes, the treatment of tense and aspect has been greatly simplified; among the issues that we will not treat is the requirement for the verb following *is* to appear in present participial form. See Butt et al. (1996) and Frank and Zaenen (2002) for discussion and analysis.

2.5 Grammatical agreement

The lexical entry for the verb *is* must specify subject agreement requirements: *is* requires a third-person singular subject. We will see in later chapters that agreement can depend on the interaction of a number of syntactic, semantic, and information-structural factors. However, in many languages, including English, grammatical agreement depends on purely syntactic factors, and can be described in simple terms, as we now show.

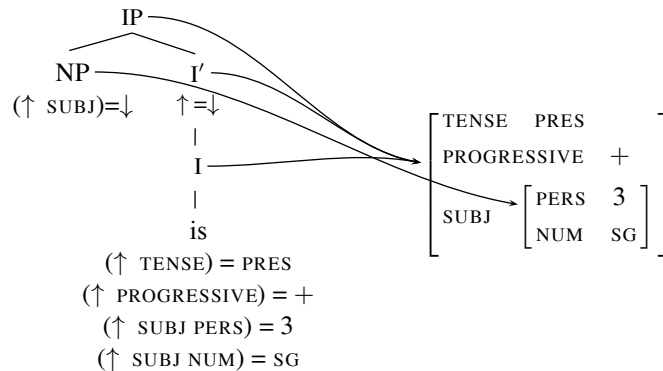
English subject/verb agreement is handled by specifications placed by the verb on features of its subject, as shown in the more complete entry for *is* in (20):

(20) Lexical entry for *is*, incorporating subject/verb agreement specifications:

is (↑ TENSE) = PRES
 (↑ PROGRESSIVE) = +
 (↑ SUBJ PERS) = 3
 (↑ SUBJ NUM) = SG

According to this entry, the subject must be compatible with the third person singular agreement features imposed by *is*. A partial tree and f-structure incorporating this lexical entry is:

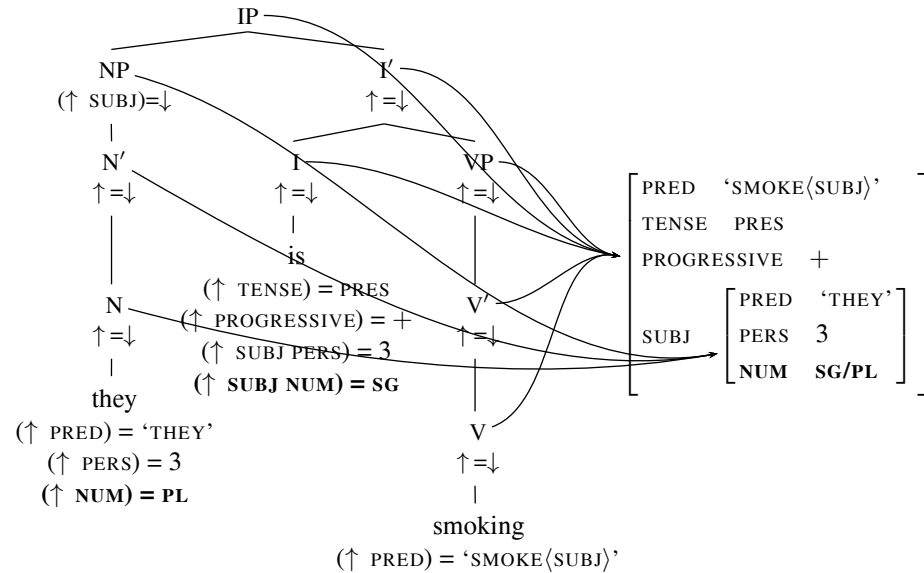
(21)



Since the person and number values specified for the subject *David* are compatible with those specified by the verb *is*, *David* is an acceptable subject of this verb. This is why the sentence *David is smoking* is acceptable, and has the c-structure and f-structure shown in (12).

In contrast, the sentence *They is smoking* is ungrammatical. This is due to incompatible specifications for the value of the NUM feature of the subject, as indicated by the clashing values SG/PL in (22): *they* has plural (PL) number, while *is* requires its subject's number to be SG.

(22) Ill-formed f-structure for **They is smoking*.



English verb agreement is particularly easy to characterise: the verb agrees with the subject, regardless of its semantic features or information-structural role. As we will see, agreement patterns in many languages are more complex and require reference to nonsyntactic levels of linguistic structure, to be presented in Chapter 4.

2.6 Agreement and pronominal incorporation

Grammatical agreement is crucially different from pronominal object incorporation, though there are clear historical relations between incorporated pronouns and agreement morphology (Givón 1976, Greenberg 1977). Bresnan and Mchombo (1987) analyse object marking in Chicheŵa as what they call **anaphoric** agreement, involving the incorporation of an object pronoun. Our concern is differential object marking as manifested in grammatical agreement morphology, not the distribution of incorporated object pronouns or clitics, and so we summarise their analysis mainly to show how anaphoric agreement is analysed in LFG and how it differs from grammatical agreement.

Chicheŵa is a subject pro-drop language: a full noun phrase subject may but need not appear. Example (23) has subject marking but no object marking. An object phrase is required, but the subject phrase may be absent:

- (23) a. njũchi zi-ná-lúm-a alenje
 bees 10Subj-Past-bite-Indic hunters
 ‘The bees bit the hunters.’
- b. zi-ná-lúm-a alenje
 10Subj-Past-bite-Indic hunters
 ‘They bit the hunters.’ (Bresnan and Mchombo 1987:744–745)

Bresnan and Mchombo propose that the verb *zi-ná-lúm-a* ‘bit’ optionally specifies a pronominal subject, and also specifies that the subject must be of noun class 10:

- (24) *zi-ná-lúm-a* ((↑ SUBJ PRED) = ‘PRO’)
 (↑ SUBJ NOUNCLASS) = 10

The equation (↑ SUBJ PRED) = ‘PRO’ is optional, and need not be used in the analysis of the sentence. When there is a full noun phrase subject, as in (23a), the pronominal contribution provided by the verb is not used, since it would clash with the specifications provided by the noun phrase subject. In this case, the verbal morphology behaves as agreement marking: the agreement features specified by the verb (here, NOUNCLASS) must match the features of the subject, just as in the English example discussed in the previous section.

When no full noun phrase subject is present, as in (23b), the pronominal contribution provided by the verb is required, since otherwise the sentence would lack a subject. In this case as well, the subject is specified as noun class 10, but additionally the verb makes a pronominal contribution associated with its subject. This is possible because of the optionality of the SUBJ PRED ‘PRO’ contribution of the verb. The functional structures for the examples in (23) are given in (25):

- (25) a. njũchi zi-ná-lúm-a alenje
 bees 10Subj-Past-bite-Indic hunters
 ‘The bees bit the hunters.’
- $$\left[\begin{array}{l} \text{PRED} \text{ ‘BITE(SUBJ,OBJ)’} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED ‘BEES’} \\ \text{NOUNCLASS 10} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{PRED ‘HUNTERS’} \\ \text{NOUNCLASS 2} \end{array} \right] \end{array} \right]$$
- b. zi-ná-lúm-a alenje
 10Subj-Past-bite-Indic hunters
 ‘They bit the hunters.’
- $$\left[\begin{array}{l} \text{PRED} \text{ ‘BITE(SUBJ,OBJ)’} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED ‘PRO’} \\ \text{NOUNCLASS 10} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{PRED ‘HUNTERS’} \\ \text{NOUNCLASS 2} \end{array} \right] \end{array} \right]$$

Chicheŵa transitive verbs may also show object marking. Bresnan and Mchombo (1987) provide compelling evidence to show that in Chicheŵa examples like (26a) and (26b), the object agreement affix is actually an incorporated pronoun, unlike the subject agreement affix, which encodes either an incorporated pronoun or grammatical agreement.

- (26) a. njũchi zi-ná-wá-luma
 bees 10Subj-Past-2Obj-bite
 ‘The bees bit them.’
- $$\left[\begin{array}{l} \text{PRED} \text{ ‘BITE(SUBJ,OBJ)’} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED ‘BEES’} \\ \text{NOUNCLASS 10} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{PRED ‘PRO’} \\ \text{NOUNCLASS 2} \end{array} \right] \end{array} \right]$$
- b. njũchi zi-ná-wá-luma alenje
 bees 10Subj-Past-2Obj-bite hunters
 ‘The bees bit them, the hunters.’
- $$\left[\begin{array}{l} \text{PRED} \text{ ‘BITE(SUBJ,OBJ)’} \\ \text{TOPIC} \left[\begin{array}{l} \text{PRED ‘HUNTERS’} \\ \text{NOUNCLASS 2} \end{array} \right] \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED ‘BEES’} \\ \text{NOUNCLASS 10} \end{array} \right] \\ \text{OBJ} \left[\begin{array}{l} \text{PRED ‘PRO’} \\ \text{NOUNCLASS 2} \end{array} \right] \end{array} \right]$$
-

When an overt noun phrase appears and is interpreted as the object, as in (26b), it is in fact a floating topic phrase, anaphorically linked to the incorporated

pronominal object.⁵ Bresnan and Mchombo provide a number of tests showing that this is the correct analysis: for example, they show that the relation between the floating topic phrase and the verb can be nonlocal, since it is the incorporated pronoun and not the full noun phrase that is the syntactic object of the verb. These tests allow us to distinguish between anaphoric agreement (an incorporated pronominal argument) and grammatical agreement (requirement for matching of features such as number, gender, and noun class). The Chicheŵa verb in the examples in (26) is associated with the following lexical specifications:

- (27) *zi-ná-wá-lum-a* (↑ PRED) = 'BITE(SUBJ,OBJ)'
 ((↑ SUBJ PRED) = 'PRO')
 (↑ SUBJ NOUNCLASS) = 10
 (↑ OBJ PRED) = 'PRO'
 (↑ OBJ NOUNCLASS) = 2

In (27), the subject and object specifications differ: the predicate PRED of the subject is optionally specified as pronominal by the verb, as indicated by the parentheses around the specification in the second line of the entry, whereas the PRED of the object is obligatorily specified. This difference gives rise to the different behaviour of the subject and object.

In this study we analyse only grammatical agreement, not anaphoric agreement, and whenever possible we provide evidence that demonstrates grammatical and not anaphoric agreement for the constructions we analyse. For further discussion of constraints on the information structure role of incorporated pronominal arguments, see Bresnan and Mchombo (1987), Culy (2000), Jaeger and Gerassimova (2002), Jaeger (2004), Bowerman (2004), and references cited there.

2.7 Casemarking

Casemarking has been a major focus of study within LFG, beginning with Nordlinger's (1998) groundbreaking theory of **constructive case**, and continuing with work by, among others, Butt and King (1999, 2003) and Butt (2008b). The constructive case view is that case signals grammatical function, and in fact "constructs" the f-structure environment in which a phrase is permitted to appear.

Nordlinger (1998) provides the following treatment of casemarking on the ergatively-marked noun *alanga-ni* 'girl-Erg' in Wambaya (Australian):

⁵We discuss the TOPIC role in functional structure and information structure in Chapter 4.

- (28) *alanga-ni* (↑ PRED) = 'GIRL'
 (↑ PERS) = 3
 (↑ GEND) = FEM
 (↑ CASE) = ERG
 (SUBJ ↑)

The specification (SUBJ ↑) is different from the equations introduced earlier, in that it uses **inside-out functional uncertainty** (Kaplan 1988) to specify the environment in which this phrase must appear. These specifications require *alanga-ni* to appear in this configuration:

- (29)
$$\left[\begin{array}{c} \text{SUBJ} \\ \left[\begin{array}{cc} \text{PRED} & \text{'GIRL'} \\ \text{PERS} & 3 \\ \text{GEND} & \text{FEM} \\ \text{CASE} & \text{ERG} \end{array} \right] \end{array} \right]$$

The first four lines in the lexical entry in (28) specify values for the features PRED, PERS, GEND, and CASE, as in the English examples discussed in Section 2.3. The specification (SUBJ ↑) in the fifth line is different: the feature SUBJ appears before the f-structure designator ↑, rather than after it, meaning that the f-structure ↑ must appear as the SUBJ within some larger f-structure.

The following Wambaya sentence satisfies the requirements in (28), since *alanga-ni* is the subject of the sentence:

- (30) *alanga-ni ngiy-a dawu darranggu*
 girl-Erg 3SgFemSubj-Past bite tree.Acc
 'The girl chopped the tree.' (Nordlinger 1998:85)

This sentence has the following c-structure tree:

- (31)
-
- ```

graph TD
 IP --> NP1[NP]
 IP --> I_prime[I']
 NP1 --> N1[N]
 N1 --> alanga_ni["alanga-ni
girl-Erg"]
 I_prime --> I[I]
 I --> ngiy_a["ngiy-a
3SgFemSubj-Past"]
 I_prime --> S[S]
 S --> V[V]
 V --> dawu["dawu
bite"]
 S --> NP2[NP]
 NP2 --> N2[N]
 N2 --> darranggu["darranggu
tree.Acc"]

```

The relevant lexical entries are the following, adapted from Nordlinger (1998:85):

- (32) *alanga-ni* (↑ PRED) = 'GIRL'  
 (↑ PERS) = 3  
 (↑ GEND) = FEM  
 (↑ CASE) = ERG  
 (SUBJ ↑)
- ngiy-a* (↑ TENSE) = PAST  
 (↑ SUBJ PERS) = 3  
 (↑ SUBJ GEND) = FEM  
 (↑ SUBJ NUM) = SG  
 (↑ OBJ PERS) = 3
- dawu* (↑ PRED) = 'CHOP(SUBJ,OBJ)'
- darranggu* (↑ PRED) = 'TREE'  
 (↑ PERS) = 3  
 (↑ CASE) = ACC  
 (OBJ ↑)

The lexical entry for *alangani* contains a specification for PRED, PERS, CASE, and GEND; it also contains the constructive case specification (SUBJ ↑), which requires *alanga-ni* to bear the grammatical function SUBJ within its clause. The auxiliary *ngiya* specifies tense as well as subject and object agreement: the subject must be third person feminine singular, and the object must be third person. The transitive verb *dawu* simply requires a SUBJ and an OBJ. The accusative noun *darranggu* specifies PRED, PERS, and CASE for its f-structure, and also contains the specification (OBJ ↑), which requires it to appear as the OBJ of its clause. These requirements are all met in the f-structure for this sentence:

- (33) 
$$\left[ \begin{array}{l} \text{PRED} \quad \text{'CHOP(SUBJ,OBJ)'} \\ \text{TENSE} \quad \text{PAST} \\ \text{SUBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'GIRL'} \\ \text{CASE} \quad \text{ERG} \\ \text{PERS} \quad \text{3} \\ \text{NUM} \quad \text{SG} \\ \text{GEND} \quad \text{FEM} \end{array} \right] \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'TREE'} \\ \text{PERS} \quad \text{3} \\ \text{CASE} \quad \text{ACC} \end{array} \right] \end{array} \right]$$

In Wambaya, the distribution of case is governed purely syntactically, marking arguments as bearing the *SUBJ* or *OBJ* role within their clause. It is clear, however, that other levels of linguistic structure also govern the distribution of case, and we now turn to a proposal to handle this within the constructive case paradigm.

## 2.8 Nonsyntactic criteria for casemarking patterns

The foregoing analyses of agreement and case depend only on syntactic features appearing within the *f*-structure: person, number, gender, and noun class features for agreement, and grammatical functions such as *SUBJ* and *OBJ* for constructive case. Nonsyntactic criteria may also be involved in grammatical marking; indeed, the main theme of our book is how information structure roles can influence grammatical marking. Here we outline one influential proposal for handling patterns in which nonsyntactic features influence casemarking.

Butt and King (2003) present a constructive analysis of case which crucially incorporates the traditional distinction between semantic case (where casemarking depends on or contributes some semantic feature) and structural case (which depends purely on grammatical function, as in the Wambaya example discussed above); see Butt (2008b) for extensive discussion of the distinction between structural and semantic case and its role in the analysis of casemarking within LFG. Butt and King provide a constructive case analysis of Georgian (Kartvelian), in which semantic case is dependent on “semantic inferences over parameters such as volitionality” (Butt and King 2003) as well as grammatical function.

Case patterns in Georgian depend on the tense and aspect of the clause; subjects of transitive aorist verbs are marked with ergative, and objects of present tense verbs are marked with accusative. Example (34) has a transitive aorist verb, and so the subject is marked with ergative case, and the object is marked nominative:

- (34) *nino-m Ceril-i daCera*  
 Nino-Erg letter-Nom wrote.3SgSubj.3Obj  
 ‘Nino wrote a letter.’ (Butt and King 2003)

Butt and King (2003) provide the following entry for the ergative case ending:

- (35) *-m(a)*      (↑ CASE) = ERG  
                   (SUBJ ↑)  
                   (EXT-ARG ↑*arg-str*)  
                   ((SUBJ ↑) TENSE-ASPECT) = AORIST

This entry specifies that the noun it is suffixed to:

- has ergative case;
- is the subject of its clause;
- is the external argument at argument structure; and
- is the subject of a clause whose tense/aspect is aorist.

The first two specifications concern the case and grammatical function of the ergative noun, and are exactly the same as for ergative nouns in Wambaya, described above.

The third specification concerns not c-structure or f-structure, but another level of structure: Butt and King use the notation  $\uparrow_{arg-str}$  to refer to the **argument structure** representation of the ergative noun, and the specification (EXT-ARG  $\uparrow_{arg-str}$ ) to require the noun to bear the argument structure role of external argument (EXT-ARG), in exactly the same way that the specification (SUBJ  $\uparrow$ ) specifies that the phrase must bear the subject role at f-structure. This means that ergative case in Georgian is semantic case, specifying not only a particular f-structure role (subject) but also a particular argument structure role (external argument). The following configuration shows the c-structure, f-structure, and argument structure that are required in Butt and King's analysis for the ergative noun *nino-m*:

|                                                      |                                                                                                                                   |                                                                          |
|------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| (36) C-structure:                                    | F-structure<br>(preliminary):                                                                                                     | Argument structure:                                                      |
| $\begin{array}{c} N \\   \\ \text{Nino} \end{array}$ | $\left[ \text{SUBJ} \left[ \begin{array}{cc} \text{PRED} & \text{'NINO'} \\ \text{CASE} & \text{ERG} \end{array} \right] \right]$ | $\left[ \text{EXT-ARG} \left[ \text{PRED} \text{'NINO'} \right] \right]$ |

The fourth specification, ((SUBJ  $\uparrow$ ) TENSE-ASPECT) = AORIST, ensures that the clause of which the ergative noun is the subject is aorist. The expression (SUBJ  $\uparrow$ ) refers to the larger f-structure within which the ergative noun is a subject; this f-structure is required to have the value AORIST for the feature TENSE-ASPECT. This results in the following requirements, taking into account all four specifications for the ergative marker:

|                                                      |                                                                                                                                                                          |                                                                          |
|------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| (37) C-structure:                                    | F-structure:                                                                                                                                                             | Argument structure:                                                      |
| $\begin{array}{c} N \\   \\ \text{Nino} \end{array}$ | $\left[ \text{SUBJ} \left[ \begin{array}{cc} \text{PRED} & \text{'NINO'} \\ \text{CASE} & \text{ERG} \\ \text{TENSE-ASPECT} & \text{AORIST} \end{array} \right] \right]$ | $\left[ \text{EXT-ARG} \left[ \text{PRED} \text{'NINO'} \right] \right]$ |

We adopt the Nordlinger/Butt and King view of constructive case and their distinction between semantic case and structural case; as we will see, we will

find it necessary to augment the treatment of semantic case to encompass information structure requirements as well as requirements at the level of semantic structure or argument structure.

## **2.9 Conclusion**

We now have our set of basic formal tools for the analysis of differential object marking in hand. Our main concern is the interaction between information structure and grammatical functions, and so in the following we will be primarily concerned with f-structure and its relation to information structure, to be defined and discussed in detail in Chapter 4. In the next chapter, we will present our basic assumptions about information structure and its relation to other linguistic levels.





## 3

# *Information structure in grammar*

Recent research on information structure addresses two main topics: the content of information structure, and the development of models of grammar which account for the interactions among information structure, syntax, and semantics. This chapter briefly deals with the first topic. It is not our goal here to evaluate theories of information structure in detail (for thorough overviews, see Vallduví 1992 and Erteschik-Shir 2007). Instead, we offer some observations on how information structure roles are defined in what we consider to be the most articulated and coherent views of information structure, with the aim of providing working definitions which can be applied in further analysis. The second question, the place of information structure in our theory of grammar, will be addressed in Chapter 4.

### **3.1 The content of information structure**

We view **exchange of information** as the main function of language. **Information structure** is the level of sentence organisation which represents how the speaker structures the utterance in context in order to facilitate information exchange. Specifically, it indicates how the propositional content of an utterance fits the addressee's state of knowledge at the time of utterance. In human communication, new information is normally added to the already existing sum of knowledge in the addressee's mind. The distinction between familiar knowledge and the informational contribution of an utterance is manifested linguistically: propositions can receive different formal expression (are "packaged") in accordance with what the speaker assesses as old or new information for the addressee.

Vallduví (1992), Vallduví and Engdahl (1996) and Erteschik-Shir (1997, 2007) view information structure as codifying information update. Following Heim (1982) and Reinhart (1982), Vallduví and Engdahl (1996) and Erteschik-Shir (2007) understand information states as file-like constructs. Files are collections of file cards corresponding to discourse referents. Each file card has a number of records that contain information about the entity it denotes; these records encode what is known about the discourse referent in question. Definite noun phrases correspond to already existing cards, while indefinite noun phrases introduce new cards. Information structure is represented metaphorically in terms of instructions for manipulating the card files. The cards can be positioned on top of the file, and their content can be altered or updated. Different types of information structuring can be viewed as different types of updating instructions. Like the grammatical relations encoded in f-structure, these instructions and the information structure units associated with them are considered universal theoretical primitives, independent of their actual linguistic realisation.

Lambrech (1994) does not employ the file card metaphor; his theory is conceptualised in terms of the mental representation of discourse referents and states of affairs that the speaker and the addressee have in their minds. As conceptual representations of states of affairs, propositions added to context are paired with the speaker's assumptions about the addressee's state of knowledge and attention at the time of an utterance. Information structure mediates between sentence meaning and form by creating a pragmatically structured proposition which reflects these assumptions. Both morphosyntactic and prosodic cues may be used to distinguish among such pragmatically structured propositions.

Lambrech emphasises that information is propositional in nature and that by communicating with the addressee, the speaker influences the addressee's mental representation of the world. This representation consists of a set of propositions, **old information**, which the addressee knows or believes at the time of speech. **New information** consists of propositions (not referents) of which the addressee does not have a mental representation.

A proposition may in principle be structured into the **pragmatic presupposition** and the **pragmatic assertion**. The pragmatic presupposition is old information: that is, a set of propositions which the speaker assumes that the addressee knows at the time of the utterance. The pragmatic assertion is new information, the proposition which the addressee is expected to learn as a result of hearing the sentence. All utterances express a pragmatic assertion in order to be informative, but the pragmatic presupposition may be absent. Obviously, pragmatic structuring is dependent on linguistic and extralinguistic environment: propositions undergo structuring according to the discourse situations

in which they appear. For this reason, we consider sentences not in isolation, but provided with a context, where possible.

### 3.2 Information structure roles

Following Lambrecht (1994), we adopt a distinction between the pragmatic presupposition and the pragmatic assertion, which forms the basis of our definitions of the units of information structure. These concepts underpin our definitions of focus and topic.

#### 3.2.1 Focus

In generative syntax and semantics, focus is often treated as a kind of operator expressing exhaustiveness (Szabolcsi 1981), identification (É. Kiss 1995), contrastiveness (Rooth 1992, Féry and Krifka 2008) or the like. This “quantificational” approach to focus contrasts with the “informational” approach we adopt here. Vallduví (1992) and Vallduví and Engdahl (1996) describe **focus** as the informative, newsy and contrary-to-expectation part of the proposition — the actual information update potential of a sentence. As such, focus is opposed to “ground”, the noninformative, known or predictable part. Focus is a relational notion in the sense that it is not the focus referent itself that is necessarily new for the addressee, but the fact that it participates in the proposition conveyed by the sentence and fills the informational gap between assertion and presupposition in a given communicative context. Focus can therefore be defined as “the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition” (Lambrecht 1994:213). The new information conveyed by a sentence (the pragmatic assertion) is the relation between the focus and the presuppositional part of the proposition.

Since the focus corresponds to an informationally unpredictable part of the proposition, it must be overtly expressed by one or several sentence elements. Some syntactic forms are explicit markers of focus; for example, *wh*-questions normally target so-called narrow focus or **argument focus**, which extends over one (or, in the case of multiple *wh*-questions, more than one) participant in the event, as in (1).

- (1) a. What is Bill eating?
- b. He is eating pizza.

The rough information structure representation of the answer is as follows:

- (2) a. pragmatic presupposition: *Bill is eating X*
- b. pragmatic assertion: *X = pizza*

c. focus: *pizza*

In contrast to narrow focus, wide focus or **predicate focus** serves to augment information about a particular referent, as in (3).

- (3) a. What is Bill doing? *or* What about Bill?  
 b. He is eating pizza in the kitchen.

The representation of (3b) is as follows:

- (4) a. pragmatic presupposition: *Bill is doing X*  
 b. pragmatic assertion: *X = eating pizza in the kitchen*  
 c. focus: *is eating pizza in the kitchen*

The focus here does not extend over a single participant as in (1b), but instead over the whole verb phrase. Wide focus corresponds to the traditional notion of **comment**. Lambrecht (1994) refers to its syntactic expression as the **focus domain**, suggesting that every constituent within this domain bears focus status.

Lambrecht (1994) also distinguishes a third type of focus structure, **sentence focus** (sometimes called presentational, news or thetic focus), in which the entire sentence is focused, as, for example, an answer to a question like ‘What happened?’. The sentence focus type corresponds to linkless structures in the terminology of Vallduví and Engdahl (1996), or to “stage topic” structures in the terminology of Erteschik-Shir (2007), discussed below. Since we are primarily interested in the status of objects as topic or focus, we will not discuss sentence focus in detail; neither will we address so-called verb/polarity/verum focus, restricted to the verb alone.

### 3.2.2 *Topic*

Following much previous research (Reinhart 1982, Gundel 1988, Lambrecht 1994, and others), we understand topic as the entity that the proposition is about. We essentially adopt the following definition from Gundel (1988:210):

An entity, E, is the topic of a sentence, S, iff in using S the speaker intends to increase the addressee’s knowledge about, request information about, or otherwise get the addressee to act with respect to E.

This definition reflects the widespread intuition that utterances normally contain some (known) elements about which the speaker wants to convey new information to the addressee.

Strawson (1964) in fact suggests that the truth of a statement must be assessed as “putative information about its topic”. According to Erteschik-Shir (2007:15), this condition implies that all sentences must have topics, since all sentences must be assigned a truth-value. This leaves “topicless” sentence-focus (thetic) sentences unaccounted for. To resolve this contradiction, Erteschik-Shir introduces the notion of “stage topic”, absent from Lambrecht’s model. Stage topics do not correspond to a referent, but instead describe a situation (time and place) about which the proposition is asserted. These spatio-temporal parameters are contextually defined, but do not necessarily have overt lexicogrammatical expression; see Erteschik-Shir (2007:16–17) for more discussion. However, since we will not discuss or analyse sentence-focus/thetic utterances in the following, the terminological difference between topicless structures and structures that involve stage topics is not relevant to us. The sentences of interest to us are those which have topic-focus structures.

Our notion of topic roughly corresponds to the concept of **link** proposed by Vallduví (1992) and Vallduví and Engdahl (1996). Informationally, link functions as a locus of knowledge update at an address in a card file: it picks out a specific card where information update is to be carried out, and therefore an entry under which a new proposition is to be classified. However, as noted by Erteschik-Shir (2007:11), link is actually defined as a command to switch to a particular existing address, so it is only involved in a change of address. In other words, the notion “link” only describes shifted topics and does not apply to continuous topics. Erteschik-Shir (2007:44) provides a description of topic in terms of the cognitive category of attention: if the attention of the addressee is drawn to a certain referent, then the addressee can be thought of as selecting the corresponding file card and placing it on top of the file. All existing file cards are potential topics in the discourse, but the actual topic is located on top of the file.

We will not treat topic in terms of attention, as we believe that this cognitive category may extend equally well to the category of focus. Since we do not use the Heimian card file metaphor either, for us topichood is defined directly on referents. Like focus, topic is a relational notion: it involves an “aboutness” relation between a referent and a proposition. This relation holds if the speaker assumes that the addressee considers a referent salient enough to be a potential locus of predication about which the assertion can be made. This means that topic inhabits the presuppositional part of the proposition: it is associated with the pragmatic presupposition of existence, and sentences with topics pragmatically presuppose that a referent is taken by the interlocutors to be the centre of current interest in the conversation. This is what Lambrecht (1994)

calls “topicality presupposition” or “relevance presupposition”. For example, the topic of the second sentence in (3) (repeated as 5) is the referent of the pronoun *he*, the individual Bill, because the sentence is construed as information about Bill. At the time the sentence is produced, the speaker assumes that the addressee takes Bill to be pragmatically important and can expect to be given more information about him.

- (5) a. What is Bill doing? *or* What about Bill?  
 b. He is eating pizza in the kitchen.

This example shows that a potential (though perhaps not universal) diagnostic for topichood is the “what-about” (or “tell-about” or “as-for”) test first discussed by Reinhart (1982) and Gundel (1988).

As should be clear from this discussion, we understand topic as a part of sentence grammar rather than a continuous discourse notion. The relationship between sentence topic and discourse topic is yet to be explored.

### 3.2.3 *Topicality and topic-worthiness*

A great deal of research has been devoted to the relationship between topicality and the referential properties of a noun phrase (see, for example, Reinhart 1982, Vallduví 1992, Lambrecht 1994 and references cited therein). An important distinction is to be made between topicality and what has been referred to as **topic-worthiness** (Comrie 2003) or **natural topicality** (Croft 1991). Topicality reflects the informational status of a referent, and has to do with the pragmatic relation that holds between the referent and the proposition. It depends on the speaker’s construal of the situation within the given communicative context, rather than on the noun phrase’s referential properties. The definition of topic relies on the speaker’s assumptions about the addressee’s state of interest with respect to a referent, and does not automatically involve the speaker’s assumptions about the addressee’s familiarity with a given referent (identifiability of a referent), expressed by the grammatical category of definiteness.

In contrast, topic-worthiness is “measured” in terms of the sorts of semantic features discussed in Chapter 1: features such as definiteness, specificity, person, animacy, and humanness. They are sometimes called “prominence features”. These features tend to characterise topics, and indeed most topics are definite and animate, the most preferred topics being first and second person pronouns denoting speech act participants. The interaction between topicality and animacy has been addressed by Givón (1983b), Dahl and Fraurud (1996), Yamamoto (1999), Leonetti (2003) and de Swart (2007), among others, and we do not have much to add. In a nutshell, these works argue that animates are more often topical than other referents because “we tend to think of the

world as organised around animate beings which perceive and act upon their inanimate environment” (Dahl and Fraurud 1996:160). That topics tend to be definite is also commonly known.

However, the correlation between topicality and topic-worthiness is imperfect. First, not all topic-worthy noun phrases are topics, but only those that are sufficiently salient or relevant (bear a “topicality presupposition”). Second, although topics tend to be discourse-old, definite, and human, indefinites and nonhuman referents are not excluded as potential topics, as long as their referents have a certain pragmatic status for the interlocutors.

The nature of the interaction between topicality and degrees of identifiability of referents has been much discussed. Lambrecht (1994:165–171) suggests the “topic accessibility scale”, according to which the most acceptable topic expressions are those whose referents are highly activated in the discourse or the situation of speech. These often correspond to an unaccented pronominal or a referential null. Next on the accessibility scale are accessible referents, which are clearly identifiable by the interlocutors and therefore correspond to a definite noun phrase. So-called “unused” referents are also identifiable, but they are inactive in the discourse and incur a higher cognitive cost when interpreted as the centre of predication. Their accessibility as topics (as well as their encoding as definite noun phrases) varies greatly among languages and types of discourse. Brand-new referents are unidentifiable for the addressee when new information is conveyed about them, and this explains why indefinite noun phrases are unlikely topics.

Yet indefinite topics are not completely excluded. If referential indefinites are pragmatically “anchored” (linked) to another identifiable entity in the consciousness of interlocutors, they can be interpreted as topics. Lambrecht calls such entities “brand-new anchored” and illustrates his point with the following contrast:

- (6) a. \*A boy is tall.  
 b. A boy in my class is tall.

In (6a) the subject is brand-new but unanchored. The sentence is unacceptable because “it is difficult to imagine a context in which it would be informative to predicate tallness on an unidentified subject referent. Such sentences violate the most elementary condition of relevance” (Lambrecht 1994:167). However, acceptability depends on pragmatic factors; in example (6b) the subject is still formally indefinite, but it corresponds to an brand-new **anchored** referent. The additional PP *in my class* restricts the unspecified set of all boys to the set of the boys in the speaker’s class, and therefore links/anchors the referent of the indefinite subject *a boy* to the speaker herself. As a member of this smaller

relevant set, the referent becomes more identifiable and more easily interpreted as topic.

“Pragmatically anchored” indefinites in Lambrecht’s sense are roughly the same as specific indefinites as defined by Enç (1986), Portner and Yabushita (2001), and other authors. This understanding of specificity, based on the idea of discourse linking, seems to be the most widely accepted (though see Farkas 1995 for a discussion of alternative views). Specificity involves a weak link to a previously established referent, where a weak link is defined in terms of a recoverable relation or a subset relation. The pragmatic link ensures that the referent of the specific indefinite expression is identifiable to the speaker (but not to the addressee). All definite phrases and some referential indefinites are specific in this sense.

Another context that makes an indefinite NP specific and topicalisable is relativisation. Erteschik-Shir (2007:8–9, 52–53) shows, based on patterns of topicalisation in Danish, that modifying relative clauses render indefinites specific. In Danish it is possible to topicalise an indefinite object if it is modified by a relative clause, but is impossible to topicalise a nonmodified indefinite object:

- (7) a. E pige som jeg mødte i går      gav jeg en god bog.  
 a girl that I met yesterday gave I a good book  
 ‘I gave a good book to a girl that I met yesterday.’
- b. \*En pige mødte jeg i går  
 a girl met I yesterday  
 ‘I met a girl yesterday.’ (Erteschik-Shir 2007:8)

In (7a), topicalisation by left dislocation is licensed by the fact that the object ‘girl’ has a specific interpretation. Erteschik-Shir argues that the relative clause indicates that the speaker has a particular referent in mind, unlike in nonspecific indefinites. The relative clause in (7a) introduces the referent of ‘girl’ and therefore makes it old with respect to the main clause and a possible candidate for topichood. In Erteschik-Shir’s model, this amounts to saying that relative modifiers cause a new card file to be opened for the indefinite noun and therefore make it a suitable topic.

We will also see in subsequent chapters that ongoing discourse can play a similar role in the topical interpretation of indefinites. In some languages an indefinite noun phrase can be interpreted as topic if the following clause (for instance, in coordinated constructions) adds more information about the respective referent. In the absence of such clause, indefinite NPs cannot be topics. The effect of the following context on topichood is still poorly understood (but see e.g. Portner and Yabushita 1998), but we will assume that it works in the same way as the relative clause in (7a).



These data indicate that specific indefinites may be interpreted as topics, although not as easily as definite noun phrases. The situation is different for nonspecific indefinites, since neither the speaker nor the addressee can identify the referent; this amounts to there being no referent at all (and no corresponding card in the card file). In other words, topical status forces a specific interpretation of an indefinite noun phrase (see Erteschik-Shir 1997, 2007, Geurts 2002, Leonetti 2003, Portner and Yabushita 2001, and references cited there for extended discussion).

It is often assumed that if a potential topic phrase has no referent, the statement cannot be evaluated as true or false (Strawson 1964, 1974; Reinhart 1982; Gundel 1988; Lambrecht 1994:154–156). Our definition of topic, as a referent which the utterance is presupposed to be about, also ensures that nonreferential expressions do not serve as topics. This requirement excludes expletive subjects and nonreferring quantified noun phrases as potential topics. However, anaphoric pronominal elements coindexed with a nonreferring phrase may be topical.

- (8) a. I am looking for a husband.  
       b. He should be rich and handsome.

In (8b) the pronoun *he* does not refer to a real world entity, but is associated with a discourse referent that can be identified from the previous context. The addressee can establish that the assertion is made about this entity and assess the truth-value of the proposition.

### 3.2.4 Secondary topic

The topic role is not necessarily unique. More than one referent can be under discussion at the time of the utterance, so that the utterance simultaneously increases the addressee's knowledge about both of them. Of course, there are cognitive limitations on the number of relevant entities under discussion in a particular communicative act. This limitation has to do with the nature of human attention: we cannot keep many entities in the centre of attention and talk about all of them at the same time. However, cross-linguistic grammatical evidence suggests that utterances can have at least two topics. Whether or not three topics can appear in the sentence is uncertain (cf. Lambrecht 1981:73–74); for our purposes it suffices to distinguish the **primary topic** and the **secondary topic** (Givón 1984a,b, Polinsky 1995, 1998, Nikolaeva 2001).

Consider the discourse in (9), from Lambrecht (1994):

- (9) a. Whatever became of John?  
       b. He married Rosa,

- c. but he didn't really love her. (Lambrecht 1994:148)

In (9b) the subject is topical, and the utterance is not assessed to be about the object referent (Rosa). This is an example of predicate-focus structure: the utterance is construed as a comment about the topical referent John. From the point of view of information structure, it can be represented as follows:

- (10) a. pragmatic presupposition: *John did X*  
 b. pragmatic assertion: *X = married Rosa*  
 c. focus: *married Rosa*

Here the object *Rosa* is part of the focus domain. In (9c) the situation is different: although it is construed primarily as information about John, it also increases the addressee's knowledge about Rosa, namely, the fact that she was not loved by her husband John. When (9c) is produced, both Rosa and John are salient, under discussion, and pragmatically linked in the consciousness of interlocutors. This pragmatic association between the two referents is established by the previous context, (9b). After (9b) is produced, the speaker can assume that the addressee is familiar with the referent Rosa and with the relation between Rosa and John ('John married Rosa'), and that the addressee can expect this relation to be commented on in further discourse. Thus, the communicative purpose of this utterance is to inform the listener about the relationship between two salient entities, John and Rosa. The new information for the addressee associated with (9c) is that John did not love Rosa. This can be represented in the following way.

- (11) a. pragmatic presupposition: *John stands in the relation X to Rosa.*  
 b. pragmatic assertion: *X = didn't really love*  
 c. focus: *didn't really love*

As Lambrecht argues, this difference in topicality is formally marked. In (9b) *Rosa* is an accented lexical noun phrase within the focus domain. In (9c) the same referent is expressed by an unaccented pronoun, as is typical of topics (Givón 1983a, Ariel 1988, Gundel et al. 1993). The same sentence can easily appear in the "what-about" context as a test for the topichood of *Rosa*:

- (12) a. And what about Rosa?  
 b. He didn't really love her.

Thus, both John and Rosa in (9c) may be characterised as topics. The two topics stand in a certain relation to each other, established before the relevant sentence is produced. This relation constitutes a part of the presupposed information associated with the sentence, while the new assertion is meant to update the addressee's knowledge about this relation. Nikolaeva (2001:26) defines secondary topic as "an entity such that the utterance is construed to be ABOUT the relationship between it and the primary topic". This definition does not explicitly reflect the fact that topics are ordered with respect to saliency: the primary topic is more pragmatically salient than the secondary topic. For example, although (9c) is construed as being about the relation between John and Rosa, John is a more salient participant.

This difference in saliency between two topical elements is captured in Vallduví's approach, which is by and large compatible with ours. Vallduví proposes a trinomial articulation of information structure. In his model, new information is termed "focus" and old information is termed "ground"; ground is further splittable into informationally more and less prominent material, **link** and **tail**. As mentioned above, our notion of topic roughly corresponds to Vallduví's "link". The second informational primitive used by Vallduví, tail, basically corresponds to what we refer to as secondary topic.<sup>1</sup> The new information conveyed by the sentence is to be recorded on the file card headed by link. Tail indicates a more specific means of adding information to the given address. It entails the presence of a particular record on the file card for link, and signifies that update is to be carried out by completing or modifying this record. This ensures that link and tail stand in a certain presupposed relation, just as was argued above for the primary and the secondary topic. The new assertion completes or modifies the tail entry, and therefore updates information about the relation between it and the link.

Consider the following Catalan example of the "link-focus-tail" structure from Vallduví (1992).

(13) a. How does the boss feel about broccoli?

b. L'amo [l'odia], el bròquil  
the.boss it.hates the broccoli

The boss HATES broccoli.

(Vallduví 1992:74)

Given the context in (13a), the interlocutors believe that the boss has some attitude toward broccoli at the time (13b) is produced: that is, the entry 'broccoli' (tail) is already present on the card for 'the boss' (link), and the relation between them is under discussion. The focus, indicated here with square

<sup>1</sup>Erteschik-Shir (2007:13) explicitly refers to both link and tail as "topic types".

brackets, substitutes for the missing material in the ‘boss/broccoli’ relation and therefore updates the information on the file card for link.

The same structure is illustrated by the following example from Erteschik-Shir (2007) (bracketing is hers).

- (14) What did John do with the dishes?  
       [he<sub>top</sub> [washed them<sub>top</sub>]<sub>foc</sub>]<sub>foc</sub> (Erteschik-Shir 2007:47–48)

Erteschik-Shir’s model has only two informational primitives, topic and focus, but topic and focus are not mutually exclusive, and subordinate structures are allowed. Erteschik-Shir analyses example (14) as containing two topics, the primary and the subordinate topic, but adds that primary topic takes precedence, in the sense that the truth-value is calculated with respect to the primary topic. Unlike our analysis, she assumes that each topic must have a focus associated with it. Therefore (14) is said to contain two superimposed focus domains. In our terms, the informational representation of (14) is as follows:

- (15) a. pragmatic presupposition: *John stands in the relation X to the dishes.*  
       b. pragmatic assertion: *X = washed*  
       c. focus: *washed*

To put it differently, both topical elements, ‘John’ and ‘dishes’, are excluded from the focus domain. The focus — the aspect in which the old and new information differ — only extends over the verb *washed*, and there is no “subordinate” focus-structure.

More research is needed on possible secondary topic contexts but, as far as we can tell, there are two frequent informational types that involve secondary topic. The first is illustrated in the examples above, where the focus extends over the predicate alone: *didn’t really love* in (9c), *hates* in (13) and *washed* in (14). These predicates are accented in English. The information structure associated with these examples is “topic - focused predicate - secondary topic”: i.e., they are instances of predicate-focus structure.

The second type is argument focus, as in the following:

- (16) a. Where did John kiss Rosa?  
       b. He kissed her in the kitchen.

(16b) updates the addressee’s knowledge about the relation between John and Rosa by adding the information that it was in the kitchen that he kissed her. This can be represented as follows:

- (17) a. pragmatic presupposition: *John kissed Rosa in X*

- b. pragmatic assertion:  $X = \textit{the kitchen}$
- c. focus: *in the kitchen*

In this example, the focus corresponds to the non-predicate phrase *in the kitchen*, but both the primary topic (John) and the secondary topic (Rosa) are expressed by non-accented pronouns, just as in (9c).

### 3.3 Conclusion

This chapter has explored our understanding of information structure and the crucial notion of topicality. The difference between topicality and topic-worthiness (or “prominence”) is of crucial importance for our analysis: topic-worthiness, accessed in terms of referential features, only determines the *likelihood* for a referent to be construed as topical, and does not automatically make it topical.

The main characteristic of topics is pragmatic saliency (“presupposition of saliency/relevance” or “topicality presupposition”, in Lambrecht’s (1994) terminology) attributed to a referent by the speaker in a given context. This of course makes the assessment of topicality in isolated written examples a rather difficult task; wherever possible, we will use question-answer pairs as well as the “what-about” tests mentioned above to elucidate information structure roles, following standard practise in information structure-related research.



## 4

# *Syntax and information structure*

Chapter 2 showed that patterns of agreement and casemarking are often definable in terms of purely syntactic criteria — verbs may obligatorily agree with their subjects, for example — but that criteria for marking may also refer to other, nonsyntactic levels of representation. LFG’s **projection architecture** (Kaplan 1987) allows reference not only to the syntactic levels of c-structure and f-structure but also to other linguistic levels and the relations among them. Our investigation centres on the role of information structure in regulating object marking patterns; here we discuss the representation and formal treatment of information structure and its relation to other levels of linguistic structure. We rely on Mycock’s (2009) insight that the traditional semantic structure of LFG, an important component of the “glue” approach to the syntax-semantics interface (Dalrymple 1999, 2001, Asudeh 2004), plays an important role in representing information structure relations.

### **4.1 Alternative views of information structure**

Various views of the representation of information structure and its relation to other grammatical modules have been proposed. Here we review some of these views, including early work within LFG.

#### *4.1.1 Tree-based representations*

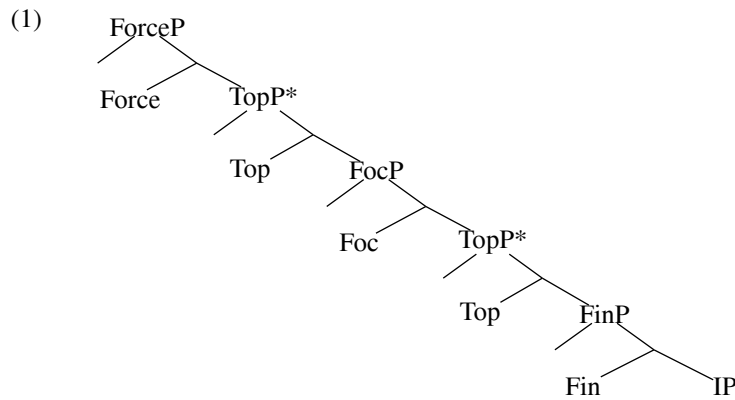
It is well known that word order and phrasal configuration are important in encoding information structure; É. Kiss (1995) discusses “discourse-configurational

languages” such as Hungarian (Uralic), which have designated phrasal positions for particular information structure roles. Within LFG, the relation between phrase structure position and information structure role has been explored in detail by, among many others, Choi (1999), Butt and King (1996, 2000), and Mycock (2006). As we discuss in Section 4.2.2 of this chapter, this work takes the standard LFG view that languages may use phrase structure position (as well as a range of other cues) to signal various kinds of grammatical information, including information structure role. Some languages (those often termed “configurational”) associate grammatical functions such as subject and object with particular phrase structure positions, and use other means — for example, prosody or morphological marking — to signal information structure. Other languages associate phrase structure positions with information structure roles (these are the “discourse configurational” languages), and use other means, often morphological, to signal grammatical function. Languages may also use a combination of the two methods, with positional and/or morphological cues signalling both syntactic and information structure roles.

Phrase structure representations play a very different role in approaches such as Principles and Parameters or the Minimalist Program, in which the phrase structure tree is the primary means of representing grammatical information. Researchers working within this paradigm often encode the information structure role of a phrase by assigning it to a particular phrase structure position (as if all languages were discourse-configurational languages, at least at an abstract level), though some have proposed the use of features or additional levels of structure to define information structure roles independent of phrasal position, as we discuss below. The second view is similar to the LFG-based approach that we adopt, since it assigns a more autonomous role to information structure, and does not assume an invariant link between phrasal position and information structure role; see Mycock (2006) for more discussion of these issues.

The influential theory of Rizzi (1997) (the “cartographic approach”), in which topic and focus appear in specifier positions of TopP and FocP phrases, exemplifies the first view. Rizzi (1997:297) provides the following schematic tree to illustrate the approach:





This structure was originally motivated by the information structure roles associated with displaced constituents in the Italian clause: any number of topics can appear before or after the single focus constituent (this is represented by the Kleene star after TopP, which indicates that TopP is a recursive category). Practitioners of the cartographic approach often assume that all topics and foci must move to the appropriate specifier positions, though possibly at a more abstract level such as LF: covert movement to these positions explains why focused elements appearing in situ can trigger effects such as weak crossover.

A number of technical problems have been raised for various versions of the cartographic view. We will not provide a comprehensive catalog of these problems, but cite only a few representative discussions: Erteschik-Shir (2007) discusses data from Italian, Hebrew, and English, Neeleman and van de Koot (2008) discuss Dutch data, and Zwart (2009) discusses Germanic data which are problematic for the approach. Some researchers working within the general Minimalist paradigm have explored alternative views, encoding information structure by means of features annotated on the phrase structure tree, or by rules which refer to general, abstract phrasal configurations.

Neeleman and van de Koot (2008) reject the cartographic view, and propose an alternative view that is still (partially) based in tree configuration, but applies more loosely. As in the LFG approach, they assume a separate level of information structure, and propose abstractly defined “mapping rules” that apply to syntactic configurations to encode/signal information structure roles, similar to the “description-by-analysis” rules that are sometimes used in LFG (Dalrymple 2001: Chapter 7). They point out that their approach does not require appeal to covert movement, though it is compatible with covert movement given some additional assumptions. If covert movement is not assumed, the approach is in broad terms compatible with an LFG-style approach, in that rules are formulated to relate syntactic structures representing the surface order of constituents to separately defined information structure representations.

Other approaches are still less tied to phrase structure, and do not rely on an association between phrasal position and information structure role. For example, Erteschik-Shir (2007) treats TOP and FOC as features which are lexically assigned to the heads of phrases, and percolate to up to the maximal phrasal projection; the features are not tied to a particular phrasal position, and may appear in various places in the tree. Similarly, Büring (2007) assumes that nodes of the phrase structure tree may be annotated with the privative features F and T, defining information structure roles of Focus and Topic, and argues against cartographic-style approaches which force covert movement in cases where there is otherwise no evidence for it.

Within the Combinatory Categorical Grammar framework, Steedman (2001) assumes a very different theory of phrase structure and its representation than is generally assumed in either the Minimalist Program or LFG. Combinatory Categorical Grammar allows the division of a clause into constituents of various types, including standard ( $X'$ -theoretic) constituents as well as sequences of words that are not constituents on either a Principles and Parameters-based or an LFG-based view. For instance, a sentence like *Anna married Manny* can have either of two constituent bracketings,  $[[Anna\ married]\ Manny]$  and  $[Anna\ [married\ Manny]]$ ; prosody determines which bracketing is chosen, as well as determining the information structure role of the constituents. Although information structure roles are in some sense aligned to phrase structure constituents, the availability of nonstandard constituents and multiple possible constituent structures for an utterance gives this approach a good deal of flexibility, and makes it crucially different from the cartographic approach. The CCG approach provides a remarkably successful theory of the relation between prosody and information structure, which differs in interesting ways from the LFG perspective. Since our focus in this work is the morphological marking of information structure roles by agreement and casemarking rather than prosody, we will not address the prosody/syntax interface issues that are central to the CCG approach, though in Section 4.2.2 below, we briefly discuss some LFG-based proposals for the treatment of prosody, its interaction with syntax and information structure, and how it helps to determine information structure role.

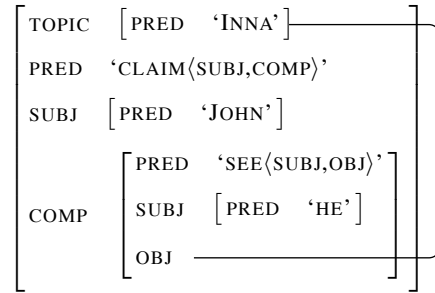
#### 4.1.2 *Early work in LFG*

There is a large body of work on information structure and information packaging in LFG, beginning with Bresnan and Mchombo (1987) and continuing with the influential work of King (1995), Butt and King (1996, 2000), and

Choi (1999); King and Zaenen (2004) and O'Connor (2006) provide a useful overview.<sup>1</sup>

Among the first LFG researchers to discuss information structure and its interactions with syntax was King (1995), in her analysis of topic, focus, and word order in Russian. King presents f-structure representations like the following:

- (2) F-structure for *Inna, John claimed that he saw (at the beach)* (King 1995:199):



Here, the f-structure for *Inna* is the TOPIC; the line connecting it to the object of the sentential complement (COMP) indicates that it plays not only the TOPIC role but also the role of the OBJ of the verb *saw*. This captures the long-distance syntactic dependency which this sentence exhibits, involving the displacement of a phrase to the beginning of the sentence (Kaplan and Zaenen 1989); the usual term for this construction is “topicalisation”, and the displaced phrase *Inna* bears the label TOPIC.<sup>2</sup>

LFG researchers commonly use the features TOPIC and FOCUS in f-structure in this way, to label displaced constituents in unbounded dependency constructions — topicalisation constructions, relative clauses, and questions — and to establish a relation between the two different roles that displaced phrases play in such constructions (TOPIC and OBJ in example (2), for example) in order to impose syntactic constraints on possibilities for displacement. When the features TOPIC and FOCUS appear at f-structure, they are taken to be **gram-**

<sup>1</sup>O'Connor (2006) refers to the level at which topic and focus are represented as d(iscourse)-structure; for him, i(nformation)-structure is a term for relations among multiple levels, including prosody and semantics. We will retain the more standard term “information structure” (or i-structure) for this level, since in some other LFG work the term “discourse structure” refers to the relations between successive utterances in a discourse, rather than to sentence-internal information packaging (King and Zaenen 2004).

<sup>2</sup>Despite the common use of the term “topicalisation” for this construction, the displaced phrase need not be an information structure topic in the sense defined in Chapter 3; in English, as discussed by Prince (1981) (see also Lambrecht 1994:31), focused constituents can also appear in this position.

**matalised** discourse functions whose synchronic function is purely syntactic, related to but different from the information structure roles of topic and focus defined and discussed in Chapter 3 (Bresnan 2001, Falk 2001).<sup>3</sup> This treatment of long-distance dependencies allows us to capture the important syntactic differences between sentences in which a constituent is displaced and those with no displaced constituents, and also encodes the historical relation between the grammaticalised discourse functions TOPIC and FOCUS in f-structure and the information structure roles to which they are related (Bresnan and Mchombo 1987). Alsina (2008) criticises this practice on the basis that it needlessly introduces (apparent) information structure distinctions into f-structure, and proposes the term OP rather than TOPIC or FOCUS for all displaced elements; we are sympathetic to this view, though for compatibility with earlier approaches, we will continue to use TOPIC and FOCUS in f-structure representations of sentences with displaced constituents rather than OP.

As noted by King (1995), there was a tendency in early LFG literature to use the f-structure labels TOPIC and FOCUS to represent not only grammaticalised (and therefore syntactic) functions, but also information structure roles as defined in Chapter 3. Although the information structure roles of topic and focus are very different from grammatical functions like subject and object, it was thought by many researchers to be convenient to co-opt the f-structure to represent these information structure roles as well as grammatical functions. When this is done, f-structure is no longer a purely syntactic representation, but instead represents a combination of information structure and grammatical structure. In their work on information structure and word order in Urdu and Turkish, Butt and King (1996) represent information structure topic and focus as f-structure features, though they note that it may be preferable to represent them at a separate level. Choi (1999) treats the information structural features  $\pm$ NEW and  $\pm$ PROM(inent) as f-structure features in her work, as we will see in Section 4.2.1 below (see also Cook and Payne 2006), though Choi asserts that information structure should be thought of as an independent grammatical representation, the nature of which is left for future work.

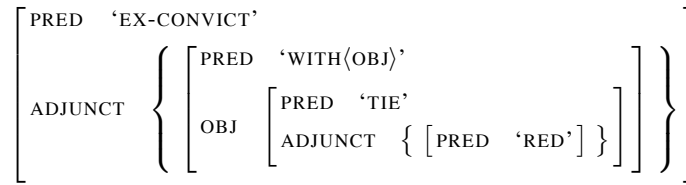
---

<sup>3</sup>Bresnan (2001) also classifies SUBJ as a grammaticalised discourse function. We do not adhere to this classification here, since we believe that the syntactic properties of the grammaticalised functions TOPIC and FOCUS are quite different from other functions appearing at f-structure: for example, verbs commonly subcategorize for the functions SUBJ, OBJ, and the oblique functions, but subcategorisation for TOPIC and FOCUS is much less common, and may not occur at all (for discussion, see Huang 1989, Her 1991, and Culy 1994). Falk (2001) refers to TOPIC and FOCUS as grammaticalised discourse functions, and classifies SUBJ, TOPIC and FOCUS as **overlay functions** (Johnson and Postal 1980), noting that each of the overlay functions “relates to its clause’s place in larger syntactic or discourse structures” (Falk 2001:59). See Chapter 5, Section 5.3 for more discussion of the relation between syntactic subject and the information structure role of topic.

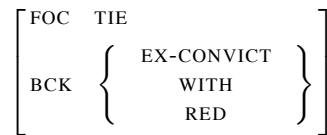
King (1997) provides a detailed critique of the practise of representing information structure roles at f-structure, particularly focusing on mismatches between f-structure and information structure: cases in which f-structure constituents are either too big or too small to adequately represent topic, focus, and other information structural constituents. Based on these and other considerations, King argues that a level of information structure separate from f-structure is necessary, and she is among the first to make an explicit proposal for a separate level of information structure (a separate **projection**, in LFG terms: Kaplan 1987, Asudeh 2006) whose parts are related by a **projection function** to the corresponding parts of other structures. King provides the abbreviated partial f-structure and i(nformation)-structure shown in (4) for example (3b):

- (3) a. Was it the ex-convict with the red SHIRT that he was warned to look out for?  
 b. No, it was an ex-convict with a red [TIE] that he was warned to look out for. (King 1997:8, citing Jackendoff 1972:232)

(4) a. F-structure (King 1997:8):



b. I-structure (King 1997:9):



We agree that it is good theoretical practise to separate out different aspects of linguistic structure and represent them separately, and we follow King (1997), Butt and King (2000), and many other LFG researchers in assuming that information structure is a separate level of representation, independent from f-structure; see O'Connor (2006) and Mycock (2006) for detailed discussion of these issues. This view also has clear similarities to the HPSG-based proposal of Engdahl and Vallduví (1996), who introduce the features FOCUS and GROUND within a structure which they call INFO-STRUCT.

## 4.2 Information structure and its role in grammar

### 4.2.1 Content of information structure: The LFG view

Choi (1999) was among the first to propose a fine-grained representation of information structure features within LFG, appealing to the features  $\pm_{\text{NEW}}$  and  $\pm_{\text{PROMINENT}}$  to encode distinctions relevant for her analysis of word order. The feature  $\pm_{\text{NEW}}$  categorises what Choi calls “discourse-newness”: focused arguments are discourse-new ( $+\text{NEW}$ ), and topic and tail (in the sense of Vallduví 1992: see Chapter 3, Section 3.2.4) are discourse-old ( $-\text{NEW}$ ). Choi classifies topic and contrastive focus as prominent or  $+\text{PROM}$ , and tail and completive focus as non-prominent or  $-\text{PROM}$ .

Butt and King (2000) adopt Choi’s classificatory features to define four information structure roles: topic is [ $-\text{NEW}$ ] and [ $+\text{PROMINENT}$ ], focus is [ $+\text{NEW}$ ] and [ $+\text{PROMINENT}$ ], completive information is [ $+\text{NEW}$ ] and [ $-\text{PROMINENT}$ ], and background information is [ $-\text{NEW}$ ] and [ $-\text{PROMINENT}$ ]. Butt and King’s classification of roles differs from Choi’s primarily in the treatment of focus: for Choi, completive focus is [ $-\text{PROMINENT}$ ], while for Butt and King it is [ $+\text{PROMINENT}$ ]. This meshes more closely with our understanding of the roles.

In Butt and King’s system, completive information is new to the addressee but, unlike focus, it is not associated with the difference between pragmatic assertion and pragmatic presupposition. According to this classification, the phrase *in the kitchen* in (5b) is a part of completive information:

- (5) a. What is Bill eating?  
 b. He is eating pizza in the kitchen.  
     TOPIC BACKGROUND FOCUS COMPLETIVE

Background information differs from topic in the following way: while topic is a pointer to the relevant information to be accessed by the addressee, background provides more detailed knowledge that may be necessary for a complete understanding of new (focused) information.

We follow Butt and King in adopting a four-way distinction in information structure roles, but also rely on the definitions of discourse functions which were presented and motivated in Chapter 3, rather than the feature-based definitions that Butt and King propose. In particular, it is important to emphasise that our secondary topic is not the same as what Butt and King (1996, 2000) refer to as background information, although in some cases these notions can overlap. There are several differences between Butt and King’s notion of background information and our notion of secondary topic. First, they differ in terms of prominence (saliency): background information is [ $-\text{PROMINENT}$ ], while secondary topic is [ $+\text{PROMINENT}$ ] (pragmatically salient, in the sense of Lambrecht 1994), just like the primary topic, although the primary topic is

more prominent than the secondary topic. We appeal to a scale of prominence in distinguishing different degrees of topichood, rather than a simple binary feature [ $\pm$ PROMINENT]. Additionally, the secondary topic appears in sentences which are construed to be about its relation with the primary topic. This implies that the proposition expressed by the sentence must involve at least two referents. There is no such requirement for background information, as is clear from example (5b), where the verb is a part of background information; indeed, some examples cited by Butt and King involve backgrounded one-place verbs and no referential NPs other than the subject.

Further, Butt and King (2000) do not distinguish between narrow and wide focus, instead associating focus with one (nonverbal) participant even in cases like (6b):

- (6) a. What is Bill doing?  
       b. He     is eating pizza in the kitchen.  
           TOPIC                                    FOCUS

On their analysis, (6b) has the same information structuring as (5b). This is primarily motivated by King's (1997) argument that there are technical difficulties with including verbs as focus or topic in information structure. However, in cases like (6b) it is difficult to determine any informational (nonsyntactic) grounds to select one particular participant for this purpose. The informational contribution of the verb in (6b) does not seem to differ from that of the nonverbal participants; the entire nonsubject portion of the sentence fills the informational gap between the speaker and the addressee. The informational role of the verb *eating* is therefore different from (5b), where it is background information.

#### 4.2.2 Linguistic encoding of information structure relations

We treat information structure as a separate, independent level of structure, containing the features TOPIC, FOCUS, BACKGROUND, and COMPLETIVE. We assume that these information structure roles are fixed by discourse context, and that assignment of these roles is subject to certain inviolable syntactic and semantic conditions discussed in Chapter 3: for example, topics must be referential, and foci must be overtly expressed. Various linguistic cues can be used to signal the information structure of an utterance to the addressee, and these cues must be consistent with the information structure roles imposed by linguistic context (for more discussion, see Mycock 2006, Erteschik-Shir 2007, Féry and Krifka 2008, and references cited there). Interestingly, such cues are often the same as those that languages use to signal grammatical functions: not only casemarking and agreement, but also word order and phrase structure

position. Choi (1999) analyses the influence of information structure roles on word order in Korean (isolate) and German (Germanic), and Butt and King (1996, 2000) provide a detailed study of the encoding of information structure roles by phrase structure position in Hindi-Urdu. For Hindi-Urdu, Butt and King show that topics appear sentence-initially, foci appear immediately before the verb, and background information is postverbal. These syntactic positions are associated with information structure roles just as grammatical functions are: by means of functional annotations on phrase structure rules, as described in Chapter 2, Section 2.4.

Prosody is also well known to play a role in the encoding of information structural relations; in this way, grammatical function encoding differs from information structural encoding, since as far as we know languages do not mark grammatical functions purely by prosodic means. Though prosody provides important indications of information structure role, it often happens that information structure roles are not unambiguously specified prosodically; Lambrecht (1994) discusses the examples in (7b) and (8b), which exhibit the same syntactic structure and the same prosody, but are associated with different information structures depending on the linguistic context:

(7) Argument-focus structure:

a. I heard your motorcycle broke down?

b. My CAR broke down.  
     FOCUS      BACKGROUND

(8) Sentence-focus structure:

a. What happened?

b. My CAR broke down. (Lambrecht 1994:223)  
     FOCUS

Linguistic cues signalling information structure role, including casemarking, agreement, word order, and prosody, must be consistent with the information structure roles imposed by linguistic context. This accounts for the unacceptability of examples such as (9), where a topical argument is associated with a prosodic contour which signals focus:

(9) a. What happened to your car?

b. \*My CAR broke down.  
     TOPIC      FOCUS



Since our primary concern is the information structure role of topic and how it is signalled by means of agreement and casemarking, we will not have much to say about the positional or prosodic encoding of information structure roles; for detailed discussion of these issues, and formal proposals within LFG for the treatment of word order and prosody that are compatible with the grammatical architecture that we assume here, see Choi (1999), Butt and King (1996, 2000), O'Connor (2006), Mycock (2006), and references cited there.

#### 4.2.3 Information structure in relation to semantics

As discussed in Chapter 3, information structure represents how the propositional content of an utterance is structured in line with the speaker's model of the addressee's state of knowledge at the time of utterance: it is concerned with utterance meaning and its "packaging" to optimise the effect of the utterance on the speaker. A formal theory of information structure must, then, involve reference to the meaning of the parts of an utterance and how they are assigned information structure roles.

Researchers in formal semantics and information structure have not agreed on the relation between truth-conditional semantics and information structure, with some researchers arguing that information structure should be represented as a completely separate module from truth-conditional semantics, and others arguing that information structure is best viewed as a means of partitioning truth-conditional meaning. We take the second view: information structure partitions sentence meaning into information structure categories, as we describe in Section 4.3 below. In this, our approach resembles "structured meaning" approaches (von Stechow 1982, Krifka 1992) in some respects, though we will see that there are important differences between our approach and theirs. We believe that it is also compatible with Lambrecht's (1994) view of information structure as the pragmatic structuring of the proposition; indeed, Lambrecht (1994:341) states that he is "not convinced that it is always possible or even useful to distinguish 'semantic meaning' from 'pragmatic meaning'".

One of the earliest proposals for structured meanings was made by von Stechow (1982), who represents utterance meanings as a list in which the first element is the topic and the remaining elements are foci. Krifka (1992) represents utterance meaning as a pair, with background as the first member and focus as the second member. Krifka (2006) assumes a three-part structured meaning for the VP 'introduced BILL to Sue', with 'BILL' in focus:

$$(10) \langle \textit{bill}, A, \lambda x.\textit{introduce}(\textit{sue}, x) \rangle \quad (\text{Krifka 2006: example 2})$$

In (10), the structured meaning is a triple, with the focus *bill* as first member, a set of alternatives to the focus *A* as second member, and the background meaning for 'introduced \_\_\_ to Sue' as third member. The set of alternatives

$A$  contains all the relevant individuals that might have been introduced to Sue, including Bill (so, for example, in the context under consideration  $A$  might be  $\{bill, fred, chris, sue, \dots\}$ ). Our approach will also assume that meanings of the parts of an utterance are separated and classified according to information structure role.

Importantly, however, we do not adopt some common assumptions that often go along with and provide motivation for structured meaning approaches. Structured meanings are often posited in the analysis of “association with focus” (Jackendoff 1972, Rooth 1985); in particular, researchers working within the structured meaning paradigm often adopt the view that meanings are structured in order to make available distinctions that are needed in the compositional semantics of so-called “focus-sensitive” operators such as *only* in examples like (11), from Krifka (2006):

- (11) a. John only introduced  $\underset{\text{FOCUS}}{\text{BILL}}$  to Sue.  
 (The only person John introduced to Sue is Bill.)
- b. John only introduced Bill to  $\underset{\text{FOCUS}}{\text{SUE}}$ .  
 (The only person John introduced to Bill is Sue.)

The truth conditions of examples (11a) and (11b) differ because of focus placement, and proponents of the structured meaning approach argue that this is best accounted for by assuming that structured meanings combine in a particular way with operators like *only*. For example, Krifka (2006) assumes that the interpretation of *only* depends on the three-part structure of the meaning of the VP *introduced BILL to Sue* given in (10). On this view, the VP adverb *only* applies to a structured meaning triple like the one in (10) to give the meaning in (12):

$$(12) \textit{only}(\langle F, A, B \rangle) = \lambda x. \forall Y \in A. [B(Y)(x) \rightarrow F = Y]$$

(Krifka 2006: example 3)

Combining the meaning for ‘only introduced BILL to Sue’ with a subject *John*, the result is:

$$(13) \forall Y \in A. [\textit{introduce}(\textit{sue})(Y)(\textit{john}) \rightarrow \textit{bill} = Y]$$

(Of the members of the alternative set  $A$ , only Bill was introduced to Sue by John.) (Krifka 2006: example 4)

A notable property of this result is that the focus/background articulation which was represented in the structured meaning in (10) is no longer present: it was,

in a sense, “consumed” by the focus-sensitive operator *only*, and the result is not a structured meaning, but what Krifka calls a “standard meaning”.

Our proposal differs in several ways. First, researchers adopting structured meanings in the analysis of operators like *only* are often concerned primarily with the distinction between focus and ground, and not with other information structure roles such as topic or the background/completive distinction. We believe that all of these information structure roles are important in the pragmatic structuring of meanings and should be represented at information structure.

Second, we do not believe that the presence of operators such as *only* results in the obliteration of the distinctions that are present in structured meanings. Schwarzschild (1997:2) makes the following observation:

Researchers have pretended that foci embedded under operators like *only* and *always* do not carry the same pragmatic import as unembedded foci. This is a surprising claim. It means that the language has an elaborate syntactic-phonological system concerned with moulding an utterance to background discourse and that this system shuts down as soon as it meets one of these operators.

We agree that this is a surprising and undesirable result. We assume that pragmatic structuring of utterance meaning is relevant for all utterances, and that this structuring does not disappear in the presence of operators like *only*; see Schwarzschild (1997) and Kadmon (2001) for more discussion of this point.

Third, we agree with many other researchers that information structure does not always provide the relevant distinctions for the interpretation of operators which have been analysed as ‘focus-sensitive’: Vallduví (1992: Chapter 7) provides useful discussion of this point, though we disagree with his conclusion that information structure must be represented as a completely separate level, unrelated to truth-conditional meaning. We believe that there are interesting and important relations between what Kadmon (2001) calls the “discourse-regulating” function of focus and its role in determining the interpretation of such operators, but we leave open the question of how exactly this should be worked out. Work by Roberts (1996), Schwarzschild (1997), Kadmon (2001), and Beaver and Clark (2008) explores this issue in depth, and shows convincingly that although focus in the sense defined in Chapter 3 often plays an important role in the interpretation of so-called focus-sensitive expressions, the relation is complex, and depends on a thorough understanding of lexical properties of operators, compositional semantics, and information structure roles. In a similar vein, Hajičová et al. (1998) explore the relation between information structure and scope, showing that while information structure often plays an important role in the determination of scope, the relation is complex and not reducible to a simple mapping. We believe that it is important to represent

information structure explicitly, and to clearly motivate and define information structure roles, in order that these connections can be thoroughly explored.

### 4.3 Our architecture

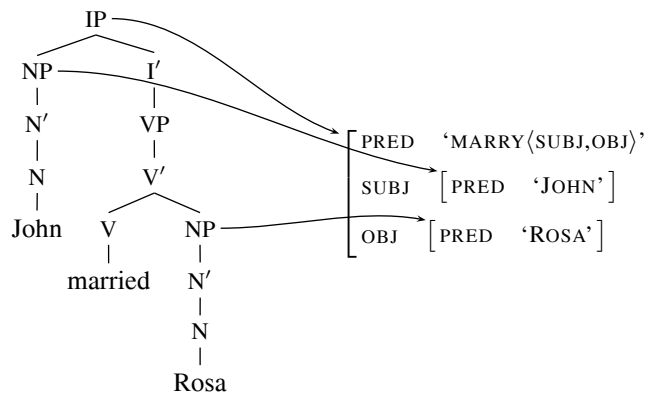
In the following, we present the formal architecture of grammar which we assume, and we show how to specify information structure relations using the formal tools of LFG. Our analysis crucially involves semantic structure in the sense familiar from work on the “glue” approach to the syntax-semantics interface (Dalrymple 1999, 2001, Asudeh 2004), and we propose a new and enriched view of the standard glue approach and its relation to information structure.

#### 4.3.1 *Glue and the syntax-semantics interface*

LFG’s theory of the syntax-semantics interface uses a resource logic, **linear logic**, to state instructions for combining meanings of the parts of an utterance to produce the meaning of the utterance as a whole. This section provides a brief overview of the standard assumptions of the glue approach, in preparation for presentation of the full formal details of our theory of the information structure module and its relation to other levels of linguistic structure. We will also provide some abbreviatory conventions for meaning expressions which are often used in the glue approach, in an effort to avoid formal overload. For a more complete introduction to glue, see Dalrymple (1999, 2001) and Asudeh (2004).

Consider the c-structure and f-structure for the sentence *John married Rosa*:

(14) John married Rosa.

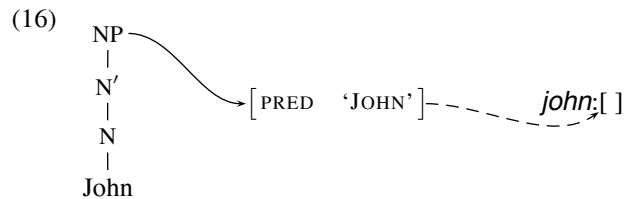


We follow the normal LFG practise of representing only the features and values of the f-structure that are relevant for current discussion, as noted in Chapter 2, leaving out (among other things) the f-structure features of tense, aspect, agreement, and case when they are not relevant for the discussion.

According to the glue approach to the syntax-semantics interface, the meaning associated with this sentence is derived via a series of instructions that can be paraphrased in the following way:

- (15) a. The word *John* contributes the meaning *john*.  
 b. The word *Rosa* contributes the meaning *rosa*.  
 c. The word *married* contributes meaning assembly instructions of the following form: When given a meaning  $x$  for my subject and a meaning  $y$  for my object, I produce a meaning  $marry(x, y)$  for my sentence.

In slightly more formal terms, glue assumes a level of semantic structure, sometimes called  $\sigma$  structure, which is related to f-structure by means of a projection function  $\sigma$  from f-structures to semantic structures. Meanings are related to expressions involving combinations of semantic structures. For the NP *John*, the following configuration is usually assumed:



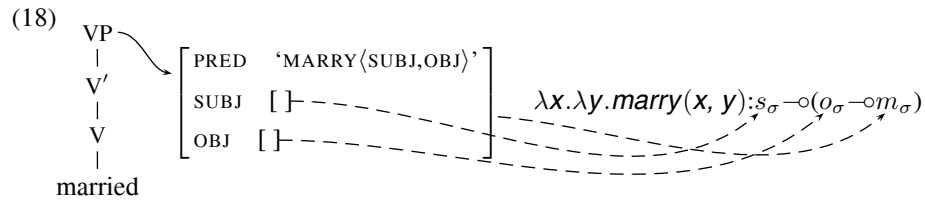
The f-structure for *John* is related to its corresponding semantic structure by the  $\sigma$  function from f-structures to semantic structures, represented by the dotted arrow. In (16), the semantic structure is represented without features or values, as is common in most glue-based literature. We will propose a set of features and values for such semantic structures below. This semantic structure is paired with the meaning for *John*, represented here simply as the term *john*. The expression *john*: [ ] consists of a meaning expression on the left side and an expression involving semantic structures on the right side, with the two sides separated by a colon: this kind of expression is called a **meaning constructor**.

The meaning constructor for *John* is represented in the lexicon as in (17), where the meaning *john* is associated with the semantic structure  $\uparrow_\sigma$  projected from the f-structure  $\uparrow$ . We use the subscript  $\sigma$  to represent the  $\sigma$  function relating f-structures to their corresponding semantic structures, which was represented as a dotted line in (16):

(17)  $john:\uparrow_\sigma$ 

The glue approach does not prescribe a particular method for the representation of meaning; any method that is adequately expressive for natural language meanings can be used. The only requirement on how meanings are expressed is that there must be an explicitly worked out way of combining meanings: in most glue-based treatments, this is function application. Dalrymple et al. (1999) discuss the use of Discourse Representation Structures (Kamp and Reyle 1993) — specifically, Lambda DRT (Bos et al. 1994) — in a glue setting, and Dalrymple et al. (1997) use intensional logic in their glue-based analysis of quantification. For simplicity, we will stick to formulas of predicate logic in the following explication.

A verb such as *married* makes a more complicated semantic contribution than a name like *John*, since it must provide instructions to combine the meanings of its subject and object to produce the meaning for the entire sentence:



The meaning of *married* is represented simply as  $\lambda x.\lambda y.marry(x, y)$ : a relation between two individuals  $x$  and  $y$  that holds if  $x$  marries  $y$ . This expression is paired with the linear logic expression  $s_\sigma - \circ (o_\sigma - \circ m_\sigma)$ , where  $s_\sigma$  is the semantic structure corresponding to the subject,  $o_\sigma$  is the semantic structure corresponding to the object, and  $m_\sigma$  is the semantic structure corresponding to the entire sentence, headed by the verb *married*. This expression involves the linear logic operator  $- \circ$ , linear implication, and can be paraphrased as follows:

(19) If I am provided with the semantic structure of my subject and then the semantic structure of my object, I produce the semantic structure of the sentence.

In the lexicon, the meaning constructor for *married* is represented as follows:

(20)  $\lambda x.\lambda y.marry(x, y):(\uparrow_{SUBJ})_\sigma - \circ ((\uparrow_{OBJ})_\sigma - \circ \uparrow_\sigma)$

The meaning constructors that are contributed by the semantically significant parts of an utterance are combined in a linear logic deduction to produce the meaning of an utterance. Linear implication on the linear logic meaning assembly side (the right side) corresponds to function application on the meaning side (the left side):

$$(21) \quad \frac{\begin{array}{l} X : f_\sigma \\ P : f_\sigma \multimap g_\sigma \end{array}}{P(X) : g_\sigma}$$

Following this rule, we can combine the meaning of the subject *John* with the meaning of the verb *married* in the following proof of the meaning of *John married*, which still requires a meaning for the object:

$$(22) \quad \frac{\text{john}:s_\sigma \quad \lambda x.\lambda y.\text{marry}(x, y):s_\sigma \multimap (o_\sigma \multimap m_\sigma)}{\lambda y.\text{marry}(\text{john}, y):o_\sigma \multimap m_\sigma}$$

We can augment the proof by providing the meaning of the object *Rosa*, producing the meaning *marry(john, rosa)* for the entire sentence, as required:

$$(23) \quad \frac{\frac{\text{john}:s_\sigma \quad \lambda x.\lambda y.\text{marry}(x, y):s_\sigma \multimap (o_\sigma \multimap m_\sigma)}{\lambda y.\text{marry}(\text{john}, y):o_\sigma \multimap m_\sigma} \quad \text{rosa}:o_\sigma}{\text{marry}(\text{john}, \text{rosa}): m_\sigma}$$

The order of combination of meaning constructors does not matter: equally, we can first combine the meaning of the object *Rosa* with the meaning of the verb *married* in a proof of the meaning of *married Rosa*, and then combine this with the meaning of the subject *John*, as shown in (24).<sup>4</sup>

$$(24) \quad \frac{\frac{\text{rosa}:o_\sigma \quad \lambda x.\lambda y.\text{marry}(x, y):s_\sigma \multimap (o_\sigma \multimap m_\sigma)}{\lambda x.\text{marry}(x, \text{rosa}):s_\sigma \multimap m_\sigma} \quad \text{john}:s_\sigma}{\text{marry}(\text{john}, \text{rosa}): m_\sigma}$$

#### 4.3.2 Information structure features and relations

Most glue language treatments assume that the meaning constructors that appear in the derivation of the meaning of an utterance are an unordered, undifferentiated collection. Our proposal differs from these: we propose to group meaning constructors according to information structure role, separating those that contribute to the focus, those that contribute to the topic, and those that contribute to background or completive information. Importantly, however, we retain the requirement of semantic completeness and coherence which other work within the glue framework assumes (Dalrymple 1999: Chapter 1; Dalrymple 2001: Chapter 9): an utterance meaning must be derivable from the

<sup>4</sup>This is an abbreviated version of the proof; in the first step we could be more explicit, and more correct, by first introducing a hypothetical meaning for the subject and combining it with the verb, then combining the resulting subject+verb meaning with the object, and then discharging the assumed meaning for the subject, and abstracting over the variable introduced as the subject meaning, to produce the verb+object meaning constructor in the second line of the proof. In the interests of simplicity, we skip these steps.

complete set of premises contributed by the parts of the utterance; the utterance meaning should not contain unsaturated expressions, in which some requirements have not been satisfied; and there should be no leftover meaning constructors which have not been used in the derivation of the meaning of the utterance.

Our proposal also differs in that we take the semantic structure to be the proper level for the representation of features relevant for the definition of the status of discourse referents, in the sense of Lambrecht (1994); in this, we follow Mycock’s (2009) insight that information structure and semantic structure are closely related. Based on Lambrecht (1994), Liao (2010) proposes a number of information structure features to represent the activation and accessibility of discourse referents: *STATUS*, whose values are *IDENTIFIABLE* and *UNIDENTIFIABLE*; *ACTV* (“activation”), whose values include *ACTIVE*, *ACCESSIBLE*, and *INACTIVE*;<sup>5</sup> and *ANCHORED*, a binary feature with a positive or negative value. Liao shows that these features are crucial in her analysis of the distribution of overt and null anaphora in Mandarin Chinese, and they also help to determine information structure role; we propose that semantic structure is the proper level for the representation of these features. Semantic features determining topic-worthiness (Chapter 3, Section 3.2.3), including animacy, humanness, definiteness, and specificity, are also best represented at semantic structure. Besides these, we introduce a semantic structure feature *DF*, whose value is specified by the linguistic context as *TOPIC*, *FOCUS*, *BACKGROUND*, or *COMPLETIVE*. This feature will be crucial in our formal analysis of the content and representation of information structure and its relation to semantic structure. In the following, we will adhere to the same policy with respect to information structure features as we do with f-structure features: though we assume that these features belong at semantic structure, if they are not directly relevant for the discussion, we will not include them in our representations.

The meaning constructors we have introduced so far, with the f-structure for the sentence *John married Rosa*, are:

$$(25) \quad m : \left[ \begin{array}{l} \text{PRED} \quad \text{'MARRY(SUBJ,OBJ)'} \\ \text{SUBJ} \quad j : \left[ \text{PRED} \quad \text{'JOHN'} \right] \\ \text{OBJ} \quad r : \left[ \text{PRED} \quad \text{'ROSA'} \right] \end{array} \right] \quad \begin{array}{l} \textit{john}:j_\sigma \\ \lambda x.\lambda y.\textit{marry}(x, y):j_\sigma \multimap (r_\sigma \multimap m_\sigma) \\ \textit{rosa}:r_\sigma \end{array}$$

In (25), the parts of the f-structure have been labelled, with the label *m* assigned to the f-structure for the entire sentence, *j* to the subject f-structure, and *r* to the object f-structure. The meaning constructors refer to  $m_\sigma$ ,  $j_\sigma$ , and

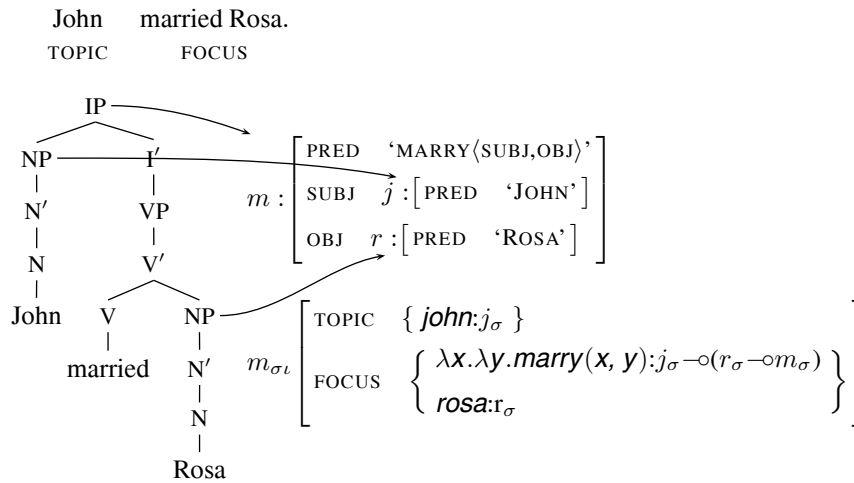
<sup>5</sup>O’Connor (2006) discusses a similar *ACTVN* feature, but treats it as binary (with a positive or negative value); Paoli (2009) makes a similar proposal for a  $\pm$ *ACTIVE* feature, which she combines with a  $\pm$ *CONTR* feature in her analysis of contrastive and new focus.



$r_\sigma$ , the semantic structures related by the  $\sigma$  function to these f-structures. Instead of representing the  $\sigma$  function from f-structures to semantic structures via a dotted arrow connecting the two structures, we will usually represent the function by means of subscripts on f-structure labels, as we did above. For example,  $f_\sigma$  is the semantic structure corresponding to  $f$ ; in other words, the function  $\sigma$  relates the f-structure  $f$  to its corresponding semantic structure  $f_\sigma$ .

We represent the syntactic, semantic, and information structural aspects of the sentence *John married Rosa* in a context in which the topic is *John* and the focus is *married Rosa* in the following way:

(26)



We introduce a new function  $\iota$  from semantic structures to information structure. An expression like  $m_{\sigma\iota}$ , which labels the information structure shown in (26), is defined in terms of the function  $\sigma \circ \iota$ , the composition of the  $\sigma$  function and the  $\iota$  function:  $m_{\sigma\iota}$  refers to the information structure which is related to the semantic structure  $m_\sigma$  by the  $\iota$  function, or equivalently the information structure which is related to the f-structure  $m$  by the composite function  $\sigma \circ \iota$ . Section 4.3.5 below provides more information about the overall architecture we assume.

We can further cut down on notational overload by introducing additional abbreviatory devices. It is standard in the glue literature to abbreviate meaning constructors by means of bold-face labels like the following:

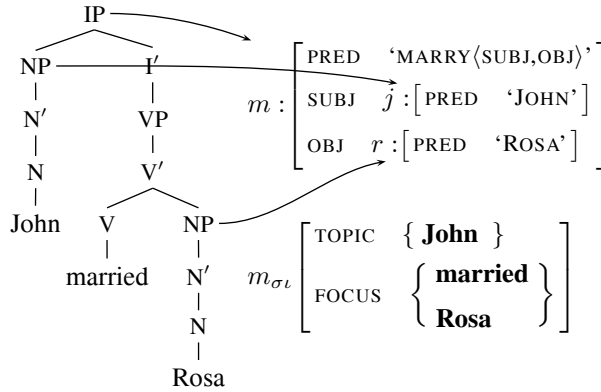
- (27) **John**                     $john:j_\sigma$   
**married**                  $\lambda x.\lambda y.marry(x, y):j_\sigma \multimap (r_\sigma \multimap m_\sigma)$   
**Rosa**                      $rosa:r_\sigma$   
**married-Rosa**          $\lambda y.marry(x, rosa):r_\sigma \multimap m_\sigma$

In (27), the meaning constructor  $john:j_\sigma$  is given the label **John**, the meaning constructor  $\lambda x.\lambda y.marry(x, y):j_\sigma \multimap (r_\sigma \multimap m_\sigma)$  is labelled **married**, and similarly for the meaning constructors labelled **Rosa** and **married-Rosa**. We can use these labels in proofs as abbreviations for the full meaning constructors: the labels are interchangeable with the meaning constructors themselves, while having the advantage of being simpler and more readable. With these abbreviations in place, we can present the proof in (24) in a more compact and readable way:

- (28) 
$$\frac{\frac{\text{married} \quad \text{Rosa}}{\text{married-Rosa}} \quad \text{John}}{\text{marry}(\text{john}, \text{rosa}): m_\sigma}$$

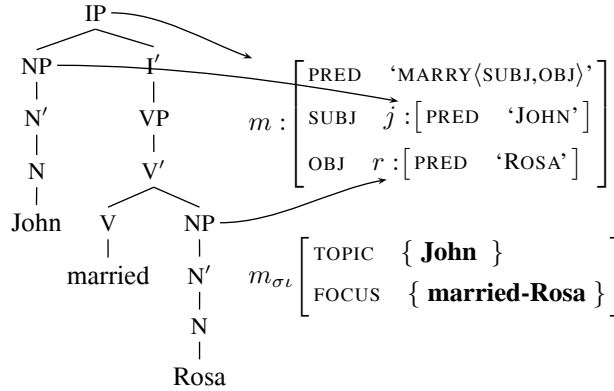
We can also use these abbreviations to present the full representation given in (26) in a simpler way:

- (29) John married Rosa.



Since the meaning constructor **married-Rosa** can be deduced via linear logic proof from the two meaning constructors **married** and **Rosa**, we could equally well represent this configuration as:

(30) John married Rosa.



This view treats TOPIC, FOCUS, BACKGROUND, and COMPLETIVE as categorising meaning contributions according to their information structure role. By virtue of its appearance in the TOPIC, FOCUS, BACKGROUND, or COMPLETIVE set, a meaning (more precisely, the meaning represented on the left-hand side of a meaning constructor) is assigned a role in affecting the context to fill the informational gap between the speaker and the addressee. In the representation above, the meaning associated with the phrase *John* is assigned a topic role, and the meaning associated with ‘married Rosa’ is assigned a focus role. Meaning constructors contributed by the various parts of an utterance are categorised according to their information structure contribution, and appear in the relevant information role category.

#### 4.3.3 Levels and equations

Our analysis depends on lexical entries like the following (again, for expository purposes we are working with the simplest possible entries, omitting much detail):

$$(31) \textit{John} \quad N \quad (\uparrow \text{PRED}) = \text{'JOHN'}$$

$$\mathbf{john} \in (\uparrow_{\sigma\iota} (\uparrow_{\sigma} \text{DF}))$$

This entry specifies that *John* is a word of category N, and is associated with a functional description consisting of two parts. The first line says simply that the node dominating the word *John* corresponds to an f-structure with the feature PRED and value ‘JOHN’, as in the f-structure in (30). The second line is crucial in achieving the desired information structure configuration:

$$(32) \mathbf{john} \in (\uparrow_{\sigma\iota} (\downarrow_{\sigma} \text{DF}))$$

This specification involves the meaning constructor  $john:j_\sigma$ , which has been abbreviated as **john**, in line with the abbreviations introduced in 27. The specification is exactly equivalent to this one:

$$(33) [john:\uparrow_\sigma] \in (\uparrow_{\sigma\iota} (\uparrow_\sigma \text{ DF}))$$

Here the meaning constructor  $john:\uparrow_\sigma$  is written inside square brackets; these are just delimiters enclosing the meaning constructor expression, and have no other significance. They could be omitted, but this would detract from readability. The entire expression can be paraphrased in the following way:

- (34) The meaning constructor  $john:j_\sigma$ , abbreviated as **john**, is a member of the set value of the discourse function signified by  $(\downarrow_\sigma \text{ DF})$  within the information structure  $\uparrow_{\sigma\iota}$ .

This constraint requires the meaning constructor for *John* to bear some information structure role.

Now, how can the proper discourse function for *John* be specified? This information must not be stated in the lexical entry for *John*, since it is not an intrinsic lexical property of *John* that it plays a particular information structure role: rather, this depends on the discourse context in which it appears on any particular occasion of its use. As discussed in Section 4.2.2, information structure roles are determined by the context of utterance, and are linguistically signalled in a number of ways: through agreement or casemarking, phrasal position, or prosody. Default information structure roles can also be associated with particular grammatical functions: for example, in many languages the subject is the default topic, as we discuss in detail in Chapter 5, Section 5.3 (see also Bresnan 2001:98), and we assume that this is the case in English as well. What is needed is a way to allow specification of these constraints, in order to determine the information structure role borne by *John*.

This is accomplished by including an feature **DF** in the semantic structure, and allowing specification of the value of **DF** as **TOPIC**, **FOCUS**, **BACKGROUND**, or **COMPLETIVE**. If the value of **DF** is specified as **TOPIC**, then the meaning constructor must be a member of the **TOPIC** set; similarly, if the value of **DF** is specified as **FOCUS**, the meaning constructor is a member of the **FOCUS** set, and similarly for **BACKGROUND** and **COMPLETIVE**. Thus, specification of a value for the semantic structure feature **DF** determines whether the meaning constructor is a member of **TOPIC**, **FOCUS**, **BACKGROUND**, or **COMPLETIVE** at information structure.<sup>6</sup>

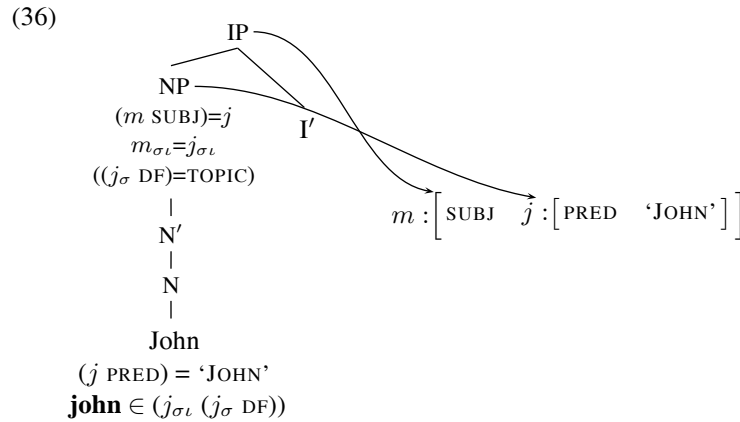
Our analysis assumes annotated phrase structure rules for English like the following:

---

<sup>6</sup>This is somewhat reminiscent of the use of the **PCASE** feature in the specification of the grammatical role of a prepositional phrase: a preposition specifies a value like **OBLGOAL** as its value for **PCASE**, and this value is then used to specify the grammatical function of the preposi-

$$(35) \text{ IP} \longrightarrow \begin{array}{c} \text{NP} \\ (\uparrow \text{ SUBJ})=\downarrow \\ \uparrow_{\sigma\iota}=\downarrow_{\sigma\iota} \\ ((\downarrow_{\sigma} \text{ DF})=\text{TOPIC}) \end{array} \quad \begin{array}{c} \text{I}' \\ \uparrow=\downarrow \end{array}$$

With *John* as the subject of the sentence, and using this rule and the lexical entry in (31), we have the following partial configuration, encompassing only the c-structure and f-structure. We have instantiated the  $\uparrow$  and  $\downarrow$  metavariables in (35) to the particular f-structure names  $m$  and  $j$  that appear in this sentence. We have also left out the  $\uparrow=\downarrow$  arrows which appear on the  $N'$  and  $N$ , which define the f-structure head relation and ensure that the NP and its head *John* correspond to the same functional structure. The arrows represent the familiar  $\phi$  function from c-structure nodes to f-structures.



We will omit the c-structure in the following exposition to avoid clutter, retaining only the functional description harvested from the annotated c-structure. Here is the full f-description:

$$(37) \begin{array}{l} (m \text{ SUBJ})=j \\ m_{\sigma\iota}=j_{\sigma\iota} \\ ((j_{\sigma} \text{ DF})=\text{TOPIC}) \\ (j \text{ PRED}) = \text{'JOHN'} \\ \mathbf{john} \in (j_{\sigma\iota} (j_{\sigma} \text{ DF})) \end{array} \quad m : [\text{SUBJ} \quad j : [\text{PRED} \quad \text{'JOHN'}]]$$

In (37):

---

tional phrase, by introducing the equation  $(\uparrow (\downarrow \text{PCASE}))=\downarrow$  at the level of the PP. If the value of  $(\downarrow \text{PCASE})$  for a particular PP is  $\text{OBLGOAL}$ , the expression  $(\uparrow (\downarrow \text{PCASE}))=\downarrow$  is exactly equivalent to the expression  $(\uparrow \text{OBLGOAL})=\downarrow$ . For detailed discussion of PCASE, see Kaplan and Bresnan (1982) and Dalrymple (2001: Chapter 6).

- The first line requires the f-structure  $j$  to be the subject of  $m$ , which is true of the f-structure shown.
- The second line requires the information structure corresponding to  $m$  and  $j$  to be the same. We assume that all members of a clause share the same information structure, and that all phrase structure rules of a language bear specifications like this one to ensure this. (We leave open the question of whether a complex, multiclausal utterance has a single information structure, or a different information structural organisation for each simple clause). The result is that specifying a particular information structure role for a meaning constructor means that it bears that information structure role within the entire clause.
- The third line provides an optional, default discourse function TOPIC for the subject; we discuss the relation between SUBJ and TOPIC in detail in Chapter 5, Section 5.3. If compatible specifications are provided by the linguistic context (see below), and the prosodic and discourse prominence features of *John* are consistent with its role as topic, the subject will be associated with the information structure role of TOPIC.
- The fourth line provides additional specification for the subject f-structure,  $j$ : it must have an feature PRED with value 'JOHN'. This is true of the f-structure shown.
- Finally, the fifth line specifies that the meaning constructor **john** must bear the role specified by  $(j_{\sigma} \text{ DF})$  at information structure.

We can simplify the final line of these constraints by assuming that the default equation  $(j_{\sigma} \text{ DF}) = \text{TOPIC}$  holds, and simplifying according to the equality  $m_{\sigma\iota} = j_{\sigma\iota}$ :

$$(38) \quad \begin{array}{l} (m \text{ SUBJ}) = j \\ (j_{\sigma} \text{ DF}) = \text{TOPIC} \\ (j \text{ PRED}) = \text{'JOHN'} \\ \mathbf{john} \in (m_{\sigma\iota} \text{ TOPIC}) \end{array} \quad \begin{array}{l} m : [\text{SUBJ } j : [\text{PRED 'JOHN'}]] \\ j_{\sigma} : [\text{DF TOPIC}] \\ m_{\sigma\iota} : [\text{TOPIC } \{\mathbf{john}\}] \end{array}$$

The equations produce the configuration shown:

- at f-structure,  $m$ 's subject is  $j$ , and  $j$ 's PRED is 'JOHN'
- the semantic structure  $j_{\sigma}$  corresponding to  $j$  has the feature DF with value TOPIC

- the value of  $j_\sigma$ 's DF appears as the feature TOPIC in the information structure for the clause,  $m_{\sigma\iota}$

In this way, meaning constructors can be specified in a particular linguistic context as bearing a particular information structure role.

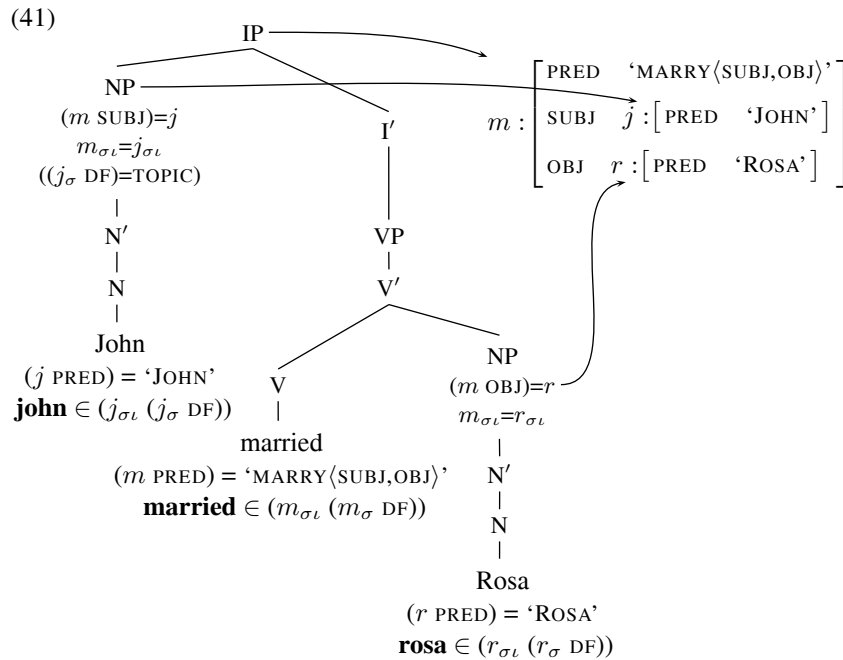
A more complete set of phrase structure rules for English is provided in (39):

$$\begin{array}{lcl}
 (39) \text{ IP} & \longrightarrow & \text{NP} \quad \text{I}' \\
 & & (\uparrow \text{SUBJ})=\downarrow \quad \uparrow=\downarrow \\
 & & \uparrow_{\sigma\iota}=\downarrow_{\sigma\iota} \\
 & & ((\uparrow_\sigma \text{DF})=\text{TOPIC}) \\
 \text{I}' & \longrightarrow & \left( \text{I} \right) \quad \text{VP} \\
 & & (\uparrow=\downarrow) \quad \uparrow=\downarrow \\
 \text{VP} & \longrightarrow & \text{V}' \\
 & & \uparrow=\downarrow \\
 \text{V}' & \longrightarrow & \text{V} \quad \left( \text{NP} \right) \\
 & & \uparrow=\downarrow \quad \left( \begin{array}{l} (\uparrow \text{OBJ})=\downarrow \\ \uparrow_{\sigma\iota}=\downarrow_{\sigma\iota} \end{array} \right)
 \end{array}$$

We do not assume a default information structure role for nonsubject constituents in English (other languages may impose a stricter relation between grammatical functions and information structure roles, as we will see in subsequent chapters) and so the annotations on these rules are the same as in standard LFG syntactic treatments, except for the specification that the object's information structure is the same as the information structure for the entire utterance (the equation  $\uparrow_{\sigma\iota}=\downarrow_{\sigma\iota}$  in the annotations on the NP daughter of V'). We also assume the lexical entries in (40) for *married* and *Rosa*:

$$\begin{array}{lcl}
 (40) \text{ married} & \text{V} & (\uparrow \text{PRED}) = \text{'MARRY(SUBJ,OBJ)'} \\
 & & [\lambda x.\lambda y.\text{marry}(x, y):(\uparrow \text{SUBJ})_\sigma \multimap ((\uparrow \text{OBJ})_\sigma \multimap \uparrow_\sigma)] \in (\uparrow_{\sigma\iota} (\uparrow_\sigma \text{DF})) \\
 \text{Rosa} & \text{N} & (\uparrow \text{PRED}) = \text{'ROSA'} \\
 & & [\text{rosa}:\uparrow_\sigma] \in (\uparrow_{\sigma\iota} (\uparrow_\sigma \text{DF}))
 \end{array}$$

We also introduce specifications for the information structure roles of *John*, *married* and *Rosa*. We assume that the discourse context identifies *John* as topical (reinforcing the default specification of the subject as topical in the phrase structure rule for IP) and *married Rosa* as in focus, and that the discourse prominence features and prosodic contour reinforce this assignment, or at least do not conflict with the assignment of these roles. We continue to use the meaning constructor abbreviations defined in (27).



Contribution from linguistic and pragmatic context:  $(j_{\sigma} \text{ DF}) = \text{TOPIC}$   
 $(m_{\sigma} \text{ DF}) = \text{FOCUS}$   
 $(r_{\sigma} \text{ DF}) = \text{FOCUS}$

We will not make an explicit formal proposal for how the equations labelled “Contribution from linguistic and pragmatic context” are introduced on the basis of the linguistic and pragmatic context. In the context under consideration for (41), the verb and object are in focus, but this may not be unambiguously signalled by casemarking, agreement, phrasal position, or prosody. A complete theory of the syntax-information structure interface would require a full specification of how the information structure roles of topic, focus, background, and completive information are determined by discourse context, how this gives rise to the equations in (41), and how all of these roles are signalled by agreement, casemarking, word order, and prosody. We hope that our proposals will form the basis of future research into these complex issues. Our aim in this book is more limited: we are interested in the information structure role of topic, how it is signalled by means of agreement and casemarking, and how the relation between topic and grammatical function is constrained; these are the issues which will be explored in the next few chapters.



We can again harvest the functional description from the annotated c-structure in (41). This time, we reorder the constraints, separating them into those that refer only to the f-structure, those that refer to semantic structure, and those that are relevant for information structure. The constraints labelled (A) specify the f-structure that is shown, and the constraints labelled (B) specify information structure roles. The constraints in (C) define the information structure for this utterance. We do not repeat the equation specifying *John* as topic, which is contextually provided as well as being optionally specified on the phrase structure rule.

$$\begin{array}{ll}
 (42) \text{ (A)} & \begin{array}{l} (m \text{ PRED}) = \text{'MARRY(SUBJ,OBJ)'} \\ (m \text{ SUBJ}) = j \\ (j \text{ PRED}) = \text{'JOHN'} \\ (m \text{ OBJ}) = r \\ (r \text{ PRED}) = \text{'ROSA'} \end{array} & m : \left[ \begin{array}{l} \text{PRED } \text{'MARRY(SUBJ,OBJ)'} \\ \text{SUBJ } j : [\text{PRED } \text{'JOHN'}] \\ \text{OBJ } r : [\text{PRED } \text{'ROSA'}] \end{array} \right] \\
 \text{(B)} & \begin{array}{l} (j_{\sigma} \text{ DF}) = \text{TOPIC} \\ (m_{\sigma} \text{ DF}) = \text{FOCUS} \\ (r_{\sigma} \text{ DF}) = \text{FOCUS} \end{array} & \begin{array}{l} j_{\sigma} : [\text{DF } \text{TOPIC}] \\ m_{\sigma} : [\text{DF } \text{FOCUS}] \\ r_{\sigma} : [\text{DF } \text{FOCUS}] \end{array} \\
 \text{(C)} & \begin{array}{l} \mathbf{john} \in (j_{\sigma\iota} (j_{\sigma} \text{ DF})) \\ \mathbf{married} \in (m_{\sigma\iota} (m_{\sigma} \text{ DF})) \\ \mathbf{rosa} \in (r_{\sigma\iota} (r_{\sigma} \text{ DF})) \\ m_{\sigma\iota} = j_{\sigma\iota} \\ m_{\sigma\iota} = r_{\sigma\iota} \end{array} & 
 \end{array}$$

We can rewrite and simplify the equations in (C) as we did above, using the equalities in (B) and in the last two lines of (C) to produce a compact description of the required information structure for this utterance:

$$(43) \text{ (C)} \quad \begin{array}{l} \mathbf{john} \in (m_{\sigma\iota} \text{ TOPIC}) \\ \mathbf{married} \in (m_{\sigma\iota} \text{ FOCUS}) \\ \mathbf{rosa} \in (m_{\sigma\iota} \text{ FOCUS}) \end{array}$$

$$m_{\sigma\iota} : \left[ \begin{array}{l} \text{TOPIC } \{ \mathbf{John} \} \\ \text{FOCUS } \left\{ \begin{array}{l} \mathbf{married} \\ \mathbf{Rosa} \end{array} \right\} \end{array} \right]$$

Further simplifying, and using the fact that **married-Rosa** can be derived by linear logic proof from **married** and **Rosa**:

$$(44) \quad m_{\sigma_L} : \left[ \begin{array}{l} \text{TOPIC} \quad \{ \mathbf{John} \} \\ \text{FOCUS} \quad \{ \mathbf{married-Rosa} \} \end{array} \right]$$

#### 4.3.4 *A short text*

With the tools provided so far, we can analyse example (9b) from Chapter 3, taken from Lambrecht (1994:148) and repeated in (45):

- (45) a. Whatever became of John?  
 b. He married Rosa,  
     TOPIC          FOCUS  
 c. but he didn't really love her.  
     TOPIC          FOCUS          TOPIC2

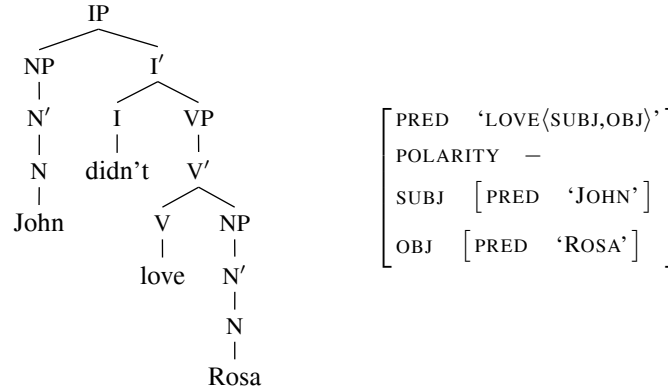
The derivation of the meaning of (45b) proceeds exactly as the derivation in the preceding section of *John married Rosa*, with *John* the topic and *married Rosa* the focus, except that the topic is realised by an unstressed pronoun rather than the name *John*.

The analysis of (45c) is more interesting, since we now have two topical phrases, the primary topic *John* and the secondary topic *Rosa*. To avoid introducing complications related to the analysis of pronouns and adverbs that are not relevant for current discussion,<sup>7</sup> we will simplify the example to *John didn't love Rosa*, assuming that the discourse context and information structure roles are the same as in (45). We then have the following simplified c-structure and f-structure:

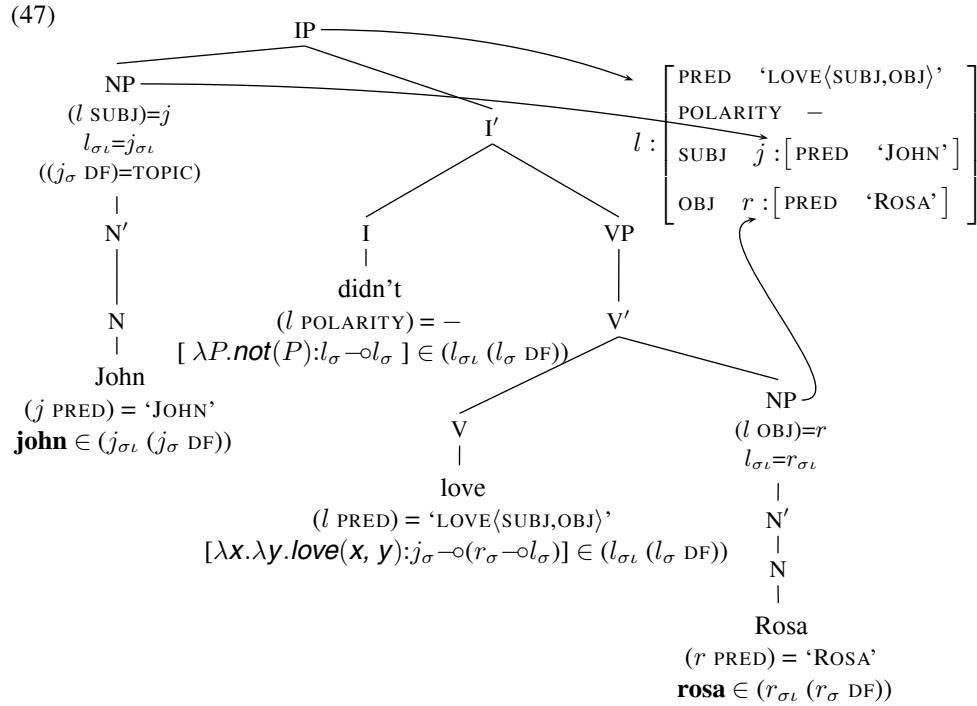
---

<sup>7</sup>For glue-based analyses of pronouns, see Dalrymple et al. (1997), Dalrymple (2001: Chapter 11), and Asudeh (2004, 2005); for a glue-based analysis of adverbial modification, see Dalrymple (2001: Chapter 10).

(46) John didn't love Rosa.



Continuing to ignore tense, aspect, and agreement features at both the syntactic and semantic levels, the c-structure annotations are largely the same as in (41) above:



We harvest these functional annotations from the c-structure, obtaining this result:

$$\begin{aligned}
 (48) \quad & (l \text{ SUBJ})=j \\
 & l_{\sigma\iota}=j_{\sigma\iota} \\
 & ((j_{\sigma} \text{ DF})=\text{TOPIC}) \\
 & (j \text{ PRED}) = \text{'JOHN'} \\
 & \mathbf{john} \in (j_{\sigma\iota} (j_{\sigma} \text{ DF})) \\
 & (l \text{ POLARITY}) = - \\
 & [\lambda P.\mathbf{not}(P):l_{\sigma}\neg l_{\sigma}] \in (l_{\sigma\iota} (l_{\sigma} \text{ DF})) \\
 & (l \text{ PRED}) = \text{'LOVE(SUBJ,OBJ)'} \\
 & [\lambda x.\lambda y.\mathbf{love}(x, y):j_{\sigma}\neg(r_{\sigma}\neg l_{\sigma})] \in (l_{\sigma\iota} (l_{\sigma} \text{ DF})) \\
 & (l \text{ OBJ})=r \\
 & l_{\sigma\iota}=r_{\sigma\iota} \\
 & (r \text{ PRED}) = \text{'ROSA'} \\
 & \mathbf{rosa} \in (r_{\sigma\iota} (r_{\sigma} \text{ DF}))
 \end{aligned}$$

We can rearrange these as before, into those labelled (A) specifying the f-structure, those labelled (B) specifying information structure roles, and those labelled (C) defining the information structure for this utterance. To the annotations derived from the c-structure above, we add the information, given in (B), that *didn't* and *love* (corresponding to the f-structure labeled *l*) are in focus, and *John* (*j*) and *Rosa* (*r*) are topical:

$$\begin{aligned}
 (49) \quad & \text{(A)} \quad (l \text{ SUBJ})=j \\
 & \quad (j \text{ PRED}) = \text{'JOHN'} \\
 & \quad (l \text{ PRED}) = \text{'LOVE(SUBJ,OBJ)'} \\
 & \quad (l \text{ OBJ})=r \\
 & \quad (r \text{ PRED}) = \text{'ROSA'} \\
 & \text{(B)} \quad (j_{\sigma} \text{ DF})=\text{TOPIC} \\
 & \quad (l_{\sigma} \text{ DF})=\text{FOCUS} \\
 & \quad (r_{\sigma} \text{ DF})=\text{TOPIC} \\
 & \text{(C)} \quad l_{\sigma\iota}=j_{\sigma\iota} \\
 & \quad \mathbf{john} \in (j_{\sigma\iota} (j_{\sigma} \text{ DF})) \\
 & \quad [\lambda P.\mathbf{not}(P):l_{\sigma}\neg l_{\sigma}] \in (l_{\sigma\iota} (l_{\sigma} \text{ DF})) \\
 & \quad [\lambda x.\lambda y.\mathbf{love}(x, y):j_{\sigma}\neg(r_{\sigma}\neg l_{\sigma})] \in (l_{\sigma\iota} (l_{\sigma} \text{ DF})) \\
 & \quad l_{\sigma\iota}=r_{\sigma\iota} \\
 & \quad \mathbf{rosa} \in (r_{\sigma\iota} (r_{\sigma} \text{ DF}))
 \end{aligned}$$

The equations in (A) characterise the f-structure shown in (47), as required, and we do not discuss them further. Simplifying the equations in (C) according to the equalities in (B) and the first and fifth lines in (C), we have:

- (50) **john**  $\in (l_{\sigma\iota} \text{ TOPIC})$   
 $[\lambda P.\text{not}(P):l_{\sigma}\text{-}o l_{\sigma}] \in (l_{\sigma\iota} \text{ FOCUS})$   
 $[\lambda x.\lambda y.\text{love}(x, y):j_{\sigma}\text{-}o(r_{\sigma}\text{-}o l_{\sigma})] \in (l_{\sigma\iota} \text{ FOCUS})$   
**rosa**  $\in (l_{\sigma\iota} \text{ TOPIC})$

This results in the following information structure:

- (51)  $l_{\sigma\iota} :$   $\left[ \begin{array}{l} \text{TOPIC} \left\{ \begin{array}{l} \mathbf{John} \\ \mathbf{Rosa} \end{array} \right\} \\ \text{FOCUS} \left\{ \begin{array}{l} \lambda x.\lambda y.\text{love}(x, y):j_{\sigma}\text{-}o(r_{\sigma}\text{-}o l_{\sigma}) \\ \lambda P.\text{not}(P):l_{\sigma}\text{-}o l_{\sigma} \end{array} \right\} \end{array} \right]$

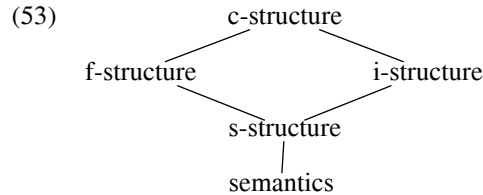
Here we have two topical elements, the primary topic *John* and the secondary topic *Rosa*. As discussed in Chapter 3, these are distinguished by prominence: we assume a scale of prominence on which the primary topic is more prominent than the secondary topic. Prominence is determined by discourse context, and can be represented by a feature or combination of features at semantic structure, where other features of discourse referents are represented. We do not specify a particular method of representing prominence here: for our purposes, it is sufficient to allow a sentence to have multiple topics, since our concern is the marking of topicality by agreement and casemarking. Indeed, in most of the languages we examine in subsequent chapters, grammatical marking does not explicitly signal secondary as opposed to primary topic, but only that the marked argument is topical.

#### 4.3.5 Information structure and its place in grammar

The content of the projections — levels of linguistic structure — which we assume, and the relations between them, are different from previous proposals within the LFG framework. We assume the overall architecture in (52), where i-structure is information structure, and s-structure is semantic structure, and the lines connecting the levels are labelled with the name of the function that establishes the relation between them:

- (52) c-structure  
 $\quad \quad \quad | \phi$   
 f-structure  
 $\quad \quad \quad | \sigma$   
 s-structure  
 $\quad \quad \quad | \iota$   
 i-structure

This follows standard LFG assumptions in assuming that the function  $\phi$  relates parts of the c-structure tree to f-structures, and the function  $\sigma$  relates f-structures to semantic structures. It differs from the approach of King (1997), who proposes the following arrangement of linguistic levels:



Similarly, Mycock (2006) proposes that information structure is directly related to both c-structure and a level of prosodic structure. King and Mycock share the view that there is a direct connection between i-structure and c-structure, while the connection between f-structure and i-structure is indirect. This view is based in King's (1997) exploration of architectural possibilities for relations between levels, discussed in Section 4.1.2 above, and her observation that f-structure constituents are often either too small or too large to define information structure roles: this is the **granularity problem**. Mycock (2006:91) raises a similar issue in her analysis of constituent questions, noting that the analysis requires reference to units that do not match f-structure constituents.

King's (1997) solution to the granularity problem was to posit a direct link between c-structure and i-structure: c-structure is more fine-grained than f-structure (for instance, V, V', VP, I', and IP are separate c-structure nodes, but all correspond to the same f-structure), and it seemed that appeal to the more fine-grained c-structure would allow definition of i-structure constituents of the proper size. Both King and Mycock posit a link between i-structure and s-structure, to ensure that all parts of the meaning of an utterance receive an information structure role.

Our approach is not susceptible to the granularity problem, which plagues approaches that treat f-structures as akin to meanings and use f-structure representations to encode information structure roles and relations. In our approach, specifying a particular information structure role for an f-structure constituent means that the **meaning constructors** contributed by the lexical head of the f-structure (or heads, if the f-structure corresponds to a functional category as well as a lexical category — recall the discussion in Chapter 2, Section 2.3) are associated with the information structure role, and does not entail that the arguments and modifiers of that head are also associated with that role. This is unlike approaches which use f-structure units to specify information structure roles, since (as King correctly points out) reference to the contents of an

f-structure also involves reference to the arguments and modifiers that appear in that f-structure. For example, in (51), the meaning constructor associated with the verb appears in the FOCUS set. This does not entail that the arguments of the verb must also appear in the FOCUS set; indeed, in (51) the subject and object of the verb are topical, not in focus.

Our approach also ensures that every meaning constructor bears some information structure role, by associating an equation of the following form with all meaning constructors:

$$(54) [\text{meaning-constructor}] \in (\uparrow_{\sigma_i} (\uparrow_{\sigma} \text{DF}))$$

According to this equation, the meaning constructor must play some role at i-structure, determined by the value of  $(\uparrow_{\sigma} \text{DF})$ . In most cases, this value is unspecified, and determined by linguistic and discourse context (although in some special cases, the value may be lexically specified: for example, question words may be intrinsically specified as FOCUS). Even though the value is unspecified, the equation requires some value to be chosen; our analysis requires all meaning constructors to be integrated into i-structure.<sup>8</sup>

An advantage of our proposal in the present setting is the direct link between f-structure and i-structure, and the concomitant ability to specify a direct relation between grammatical function and information structure role; as we will see in subsequent chapters, such specifications are important in the grammar and structure of many languages.

#### 4.4 Conclusion

We have presented a formal theory of information structure and its place in the overall architecture of Lexical Functional Grammar. Information structure is intimately related to semantic structure, which comports well with “structured meaning” approaches as well as with Lambrecht’s (1994) theory. A pleasant feature of our theory is that it allows for a simple specification of the information structure role of an argument, by providing a value for the DF feature within its semantic structure. In the following chapters, we will show how agreement and casemarking can affect the specification of DF as TOPIC for a range of arguments in the clause.

---

<sup>8</sup>In this way, our proposal differs from analyses such as Kwon and Zribi-Hertz (2008), who assume that certain unmarked subjects and objects in Korean are entirely excluded from information structure; we assume that all of the meaning contributions of a sentence play a role at information structure. If it is found that certain unmarked or incorporated elements do not bear one of the four information structure roles that we assume (TOPIC, FOCUS, BACKGROUND, or COMPLETIVE), it may be necessary to propose an additional information structure role for extremely backgrounded meaning contributions.





## *Topicality and grammatical marking*

In a number of languages, topicality affects how a sentence element is grammatically **marked**. Topical marking (casemarking and agreement) is restricted to topical **subjects** in some languages, but can be associated with any of several grammatical functions in other languages. In this chapter, we explore connections involving grammatical function, grammatical marking, and topicality, demonstrating correlations between grammatical marking and topic-hood as well as the common association between subjects and primary topics. We will discuss topicality as it affects grammatical marking of arguments within a single clause, although topicality has also been shown to be important for cross-clausal long-distance agreement, as Polinsky and Potsdam (2001) demonstrate in their analysis of Tsez.

### **5.1 Topical marking for different grammatical functions**

Languages often single out topical arguments for grammatical marking. For example, Bossong (1989) discusses Ayacucho Quechua (Quechuan), a language in which the topic marker *qa* can combine with any case marker which signals the syntactic role of clause-level NPs. In (1a), the subject shows topical marking. In (1b), the same marking is found on a locative element, and the subject is detopicalised and syntactically inverted.

- (1) a. *cheqa-paq-mi nina-para-qa chaya-ska-sqa*  
 truth-for-Foc fire-rain-Top arrive-still-Past.3Sg  
 ‘Really, the fire-rain is still falling.’ (Bossong 1989:45)
- b. *kay-pi-qa sumaq-ta-m wiña-n kawsay-kuna*  
 this-Loc-Top good-Adv-Foc grow-3Sg crop-Pl  
 ‘Here, the crops are growing well.’ (Bossong 1989:44)

Similarly, the Japanese topic marker *wa* occurs on NPs with different grammatical functions. The topic marker replaces the subject and the object markers, but can appear with some other case affixes/postpositions, as shown in (2–3):

- (2) a. *uti no kodomo ga koukou ni hairu*  
 we Gen child Nom high.school to enter  
 ‘Our child will enter high school.’ (Fry 2001:138)
- b. *uti no kodomo wa koukou ni hairu*  
 we Gen child Top high.school to enter  
 ‘As for our child, she will enter high school.’ (Fry 2001:139)
- (3) a. [*nihongo ni wa yama toka kawa toka itiminzi de*  
 Japanese in Top mountain etc. river etc. straight.line Cop  
 ‘In Japanese, ‘mountain’, ‘river’, etc. are (written with) straight lines.’  
 (Fry 2001:140)
- b. [*watasi no toko no ie kara wa kanari tooi no*  
 I Gen place Gen house from Top pretty far Focus  
 ‘From my house, it’s pretty far.’ (Fry 2001:140)

Given the typical association between topic and subject — and more generally, between topics and grammatical functions that are high on the grammatical function hierarchy — we expect topical marking to appear on subjects in the majority of cases, even when it is also permitted on nonsubjects. This is confirmed by Fry (2001), who shows that 70% of *wa*-marked nouns in a corpus of spoken Japanese are subjects, and less than 4% bear functions other than subject or object.

Agreement can work in a similar way, and indeed, Comrie (2003) and others have argued that agreeing elements are highly topical. Morimoto (2009) analyses the so-called subject-object reversal construction in the Bantu languages Kinyarwanda and Kirundi, showing that the standard term for this construction is misleading: in fact, no grammatical function change is involved. Instead,

she provides evidence that the construction involves topical agreement with a pragmatically prominent argument, which may be either a subject or an object. Kinyarwanda verb agreement in noun class is expressed by a prefix, and the agreement controller is normally the argument in immediately preverbal position. In example (4a), the verb shows noun class agreement with the topical subject 'boy'. In contrast, in (4b) 'book' is topical, and the subject is in focus. Agreement in this case is with the object.

- (4) a. umuhuungu a-ra-som-a      igitabo  
       1.boy          1-Pres-read-Asp 7.book  
       'The boy is reading the book.'                         (Morimoto 2009:201)
- b. igitabo ki-som-a      umuhuungu  
       7.book 7-read-Asp 1.boy  
       'The BOY (focus) is reading the book (topic).'
- (Morimoto 2009:201)

Morimoto supports her analysis by demonstrating that the preverbal argument in (4b) ('book') does not have subject status; her evidence comes from syntactic tests such as relativisation, gapping in coordinate structures, and subject-to-subject raising. Further, the postverbal argument in the reversal construction does not display the usual object properties: it does not have to be located in the canonical immediately postverbal object position, it does not trigger object agreement, and it can be neither relativised nor pronominalised. Morimoto concludes that the postverbal argument in the reversal construction is in fact the subject, and that agreement in Kinyarwanda and Kirundi is controlled by the preverbal topic, independently of its grammatical function (subject or object).

Aleut (Eskimo-Aleut) verb agreement also depends on information structure role. According to Golovko (2009), the verb in Aleut always agrees with a topical argument, which may be the subject, the object, the possessor of the subject or object, or an oblique goal. Aleut has two paradigms of person-number agreement with the topic, which Golovko calls primary agreement and secondary agreement. Primary agreement is with a topical subject or a topical possessor of the subject. Secondary agreement is used if a topicalised argument is an object, the possessor of the object, or in some cases the oblique goal argument of a verb like 'give'.<sup>1</sup> For example, in (5a), agreement with the topical subject is expressed via the primary paradigm. In (5b), the verb shows third person plural agreement, although the subject ('teacher') is singular. Here, agreement is with the fronted topical object, expressed via

<sup>1</sup>There are other complications in the Aleut agreement system which we do not discuss here; see Golovko (2009) for a detailed description.

the secondary paradigm; the nontopicalised subject argument takes the special “relative” form, which also marks possessors.<sup>2</sup>

- (5) a. ang'ag'ina-s suna-x' ukux'ta-ku-s  
 person-PI boat-Sg see-NonFut-3PI  
 ‘The people see the boat.’
- b. anik'du-s achixana-m achixa-ku-n'is  
 child-PI teacher-Rel.Sg teach-NonFut-3PI  
 ‘The children, the teacher teaches them.’ (Golovko 2009)

Possessor topics cannot be overtly expressed within the same clause, and must be recoverable from the preceding context. However, Golovko argues that the possessor of the subject in (6) is topical, although it is not expressed by an independent element; the verb shows singular agreement with the possessor of the subject, rather than plural agreement with the subject:

- (6) âa-n'is asxinu-un kidu-ku-x'  
 son-3Sg.Pl daughter-2Sg help-NonFut-3Sg  
 ‘His/her sons help your daughter.’ (Golovko 2009)

In contrast, overt possessors cannot be analysed as topical, and cannot control agreement:

- (7) ajaga-m âa-n'is asxinu-un kidu-ku-s /  
 woman-Rel.Sg son-3Sg.Pl daughter-2Sg help-NonFut-3PI  
 \*kidu-ku-x'  
 help-NonFut-3Sg  
 ‘The woman’s sons help your daughter.’ (Golovko 2009)

There is no independent evidence that the agreeing possessor in example (6) plays a role as argument of the main predicate; rather, agreement is with the possessor topic.

The verb may also show agreement with a goal topic of a verb like ‘give’:

- (8) a. âax' asxinu-m n'aan kanfiita-x' ag'i-ku-x'  
 boy-Sg girl-Rel.Sg to.3Sg candy-Sg give-Nonfut-3Sg  
 ‘The boy gave the girl the candy.’
- b. asxinu-x' âa-m kanfiita-x' n'aan ag'i-ku-u  
 girl-Sg boy-Rel.Sg candy-Sg to.3Sg give-Nonfut-3Sg  
 ‘The girl, the boy gave her a candy.’

<sup>2</sup>We have transliterated Golovko’s examples from Cyrillic.

In (8a), the subject is topical, and the verb shows primary third person agreement. In (8b), the goal argument ‘girl’ is topicalised and fronted, and the verb shows secondary third person agreement with the goal.

These examples illustrate that grammatical marking of primary topicality need not be restricted to one grammatical function, but may target various functions including the subject. In Ayacucho Quechua, Japanese and Korean, topical markers can cooccur with nearly every verbal dependent. Some languages display additional grammatical constraints on marking: in Kinyarwanda and Kirundi, topical agreement is possible only with the subject and the object, the two grammatical functions that are highest on the grammatical function hierarchy. Aleut primary agreement is controlled by the subject or its possessor, while topicalisation of other sentence elements may be expressed by secondary agreement. These data confirm that grammatical marking (agreement or case-marking) may target (primary) topics, and may sometimes be constrained by additional syntactic and semantic factors.

The formal analysis of topical marking in these languages is straightforward, given the theory of information structure and its relation to syntax presented in Chapter 4. The Quechua topic marker *qa* and the Japanese topic marker *wa* are associated with the constraint in (9):

(9) Topic marking (any grammatical function):

$$(\uparrow_{\sigma} \text{DF}) = \text{TOPIC}$$

This constraint ensures that the argument bearing topic marking is associated with the information structure role of TOPIC, in accordance with the theory presented in Chapter 4, Section 4.3. It does not specify a particular grammatical function for the argument it is attached to, since topic marking in these languages is syntactically unconstrained: it can be associated not only with the subject, but with other grammatical functions as well.

Topical agreement may be similarly underconstrained, and may be associated not only with the subject, but with other grammatical functions as well. For Kinyarwanda and Kirundi, topical agreement marking is controlled by either the subject or the object of the verb. The constraints associated with a verb with Class 7 agreement morphology are:

(10) Class 1 agreement with topical subjects or objects, Kinyarwanda/Kirundi:

$$(\uparrow \{ \text{SUBJ} | \text{OBJ} \}) = \%t$$

$$(\%t \text{ NOUNCLASS}) = 1$$

$$(\%t_{\sigma} \text{DF}) = \text{TOPIC}$$

This specification uses a **local name**, which is used when constraints are placed on an f-structure whose grammatical function is uncertain or undetermined; see

Dalrymple (2001:146–148) for definition and discussion. Local names always begin with a percent sign (%). The first line of the specifications defines the local name %t as referring either to the subject or to the object of the agreeing verb. The second line requires noun class 1 for %t, and the third line requires %t to bear the information structure role of topic.

In Aleut, topical agreement is controlled by the subject, the object, the oblique goal argument of a verb such as ‘give’, or a possessor of the subject or object. Third person topical primary agreement involves the following specifications:

- (11) Primary agreement with topical subjects or possessors of the subject, Aleut:

$$\begin{aligned} (\uparrow \text{SUBJ (POSS)}) &= \%t \\ (\%t \text{ PERS}) &= 3 \\ (\%t_{\sigma} \text{ DF}) &= \text{TOPIC} \end{aligned}$$

This expression also uses the local name %t to define the controller of agreement and to specify it as topical. Parentheses around POSS allow the topical argument to be the possessor of the subject: the path specifying the agreement controller can be either SUBJ or SUBJ POSS. First and second person primary topical agreement is defined similarly, except that the controller of agreement is specified as first or second person rather than third person. Third person topical secondary agreement is with the object, its possessor, or an oblique goal argument:

- (12) Secondary agreement with topical objects, possessors of the object, or oblique goals, Aleut:

$$\begin{aligned} (\uparrow \{ \text{OBJ (POSS)} \mid \text{OBLGOAL} \}) &= \%t \\ (\%t \text{ PERS}) &= 3 \\ (\%t_{\sigma} \text{ DF}) &= \text{TOPIC} \end{aligned}$$

Constraints for first and second person verbs are similar, except for the specification of the person value as 1 or 2 rather than 3.

## 5.2 Grammatical encoding of topical subjects

In some languages, subjects which are also topics are grammatically marked; this is **differential subject marking**, or subject marking which depends on topicality. Topical subjects often receive special casemarking, and also tend to trigger more agreement than nontopical or focused subjects (Comrie 2003, Siewierska 2004:159ff, Corbett 2006:197-204). Conversely, Lambrecht and Polinsky (1997) and Lambrecht (2000) argue that in sentence-focus construc-

tions where both the subject and the predicate are in focus, detopicalisation of the subject may be accompanied by suspended subject-verb agreement.

Somali (Semitic) provides good evidence for these generalisations, as shown by Saeed (1984, 1987). There are two subject forms in Somali, the absolutive case and the “subject” (nominative) case. The absolutive is the basic citation form and the case that marks some subjects, objects and obliques. The nominative is derived from the absolutive by tonal alternations and sometimes by adding a final vowel. Crucially, the nominative is found only on topical subjects, while nontopical or focused subjects must stand in the absolutive form. Furthermore, with focused subjects the Somali verb must be in the so-called “relative” form, a paradigm in which agreement is reduced: the relative paradigm includes only three distinct forms, in contrast to the five or six forms typically found in paradigms used with topical subjects.

For example, in (13a) and (14a) the subjects are topical and nominative. The particles *wuu* and *wày* encode declarative marking and positive polarity, and both the verb and the declarative particle express agreement with the subject. In contrast, (13b) and (14b) contain the focus marker *ayàa*, indicating that the subject is focused; the verb is in reduced relative form, and there is reduced verb agreement.

- (13) a. *nínku wuu imánayaa*  
 man.Nom Decl.3Sg.Masc come.Pres.Prog.3SgMasc  
 ‘The man is coming.’ (Saeed 1987:216)
- b. *nínka ayàa imánayá*  
 man.Abs Foc come.Rel.1Sg/2Sg/3SgMasc/2Pl/3Pl  
 ‘The MAN is coming.’ (Saeed 1987:216)
- (14) a. *Nimánkii wày keeneen*  
 men.Nom Decl.3Pl bring.Past.3Pl  
 ‘The men brought (it).’ (Saeed 1987:217)
- b. *Nimánki ayàa keenáy*  
 men.Abs Foc bring.Rel.1Sg/2Sg/3SgMasc/2Pl/3Pl  
 ‘The MAN brought (it).’ (Saeed 1987:217)

Other languages pattern similarly. Maslova (2003a) shows that in Kolyma Yukaghir (isolate), topical subjects trigger agreement in number (singular or plural) and person (first, second and third). Thus, the verbal paradigm for topical subjects consists of six forms. When the (intransitive) subject bears focus marking, the verb exhibits reduced agreement, distinguishing only third person

plural from the rest. Example (15) illustrates this: the subject is focused, as indicated by focus marking, and the verb shows reduced agreement:<sup>3</sup>

- (15) o, nahā omo-s'e šoromo-k kel-u-l  
 Interj very good-Attr person-Foc come-EpentheticVowel-SubjFoc  
 'Wow, a very good person has come.' (Maslova 2003a:91)

In contrast, the subject in example (16b) is topical, and the verb shows full agreement:

- (16) a. The lake-king was delighted and let all the fishes go. He sent them into the river along watercourses.  
 b. tamun jelāt tude+sam kewe-s' čobul lajin  
 that after he+self go-Perf.Intr.3Sg sea Dir  
 'And then he went away into the sea himself.' (Maslova 2003a:572)

Maslova (2003a) argues that focused and topical subjects in Kolyma Yukaghir have the same syntactic properties, despite the fact that they bear different casemarking and trigger different agreement. It is the information structure role of the subject that determines its casemarking as well as which agreement paradigm is used.

Formally, we analyse topical subject marking in a way similar to the analysis presented above, for marking of any topical element; the main difference is the additional requirement that the marked argument must be a subject. Nominative/topical casemarking in Somali is associated with the constraints in (17):

- (17) Casemarking on topical subjects:

(SUBJ ↑)  
 (↑<sub>σ</sub> DF) = TOPIC

The first of these two specifications, (SUBJ ↑), requires the argument bearing nominative/topical case (represented by the f-structure metavariable ↑) to be the subject of its clause, in line with the constructive case approach discussed in Chapter 2, Section 2.7. The second line specifies that the casemarked argument's discourse function is TOPIC. Together, these two specifications ensure that the casemarked argument is the SUBJ at f-structure, and that its semantic contribution is associated with the information structure role of TOPIC.

Agreement as a marker of topicality works similarly. Each member of the Yukaghir full agreement paradigm is associated with the following specification:

<sup>3</sup>Maslova glosses the focus ending on the focused subject as Pred, for Predicative Case. We have glossed it as Foc, for Focus marking.



(18) Agreement with topical subjects:

$$((\uparrow \text{SUBJ})_{\sigma} \text{DF}) = \text{TOPIC}$$

Here  $(\uparrow \text{SUBJ})$  is the verb's subject, and  $(\uparrow \text{SUBJ})_{\sigma}$  is the subject's semantic structure. Specifying the value **TOPIC** for the feature **DF** within the subject's semantic structure ensures that the semantic contribution of the **SUBJ** is associated with the **TOPIC** information structure role. The reduced agreement paradigm does not mark the subject as topical, and is not associated with this constraint.

### 5.3 Subjects and topichood

Chapter 3 showed that topicality is defined on referents rather than linguistic expressions, so a topic is not *a priori* constrained to be encoded as a particular grammatical function. However, there are certain cross-linguistic tendencies governing the syntactic expression of topics. We understand these as specifying the preferred mapping (alignment) between information structure and syntax: even though sentence form is not fully determined by function, some of its aspects may be motivated by function, including information structure role.

We have seen that there is a strong association between topic marking and subjecthood. This is not arbitrary: the correlation between (primary) topic and subject has been much discussed, at least since Hockett (1958) and Keenan (1976) (see Givón 1976, Comrie 1989, Bossong 1989, Yamamoto 1999, and Erteschik-Shir 2007, among others). Givón (1976) suggests that subjects are grammaticalised primary topics, and Lambrecht (1994:132–137) argues that subjects are unmarked topics. Given Lambrecht's typology of focus structures, this implies that the topic-comment articulation (wide or predicate focus) is unmarked information structuring. Predicate focus utterances serve to augment the addressee's information about a referent under discussion. This is communicatively more common than identifying an argument in an open proposition (narrow focus) or reporting a new and unexpected event (sentence focus). The subject-predicate structure, then, iconically reflects the unmarked topic-comment information structure.

More generally, this reflects a tendency toward isomorphism between pragmatic (informational), semantic, and syntactic prominence. On one hand, the topic referent is cognitively salient because it is characterised by an "aboutness" relation to the proposition (see Chapter 3, Section 3.2.2). On the other hand, the subject normally corresponds to a participant in the described situation that has the most prominent semantic role and is often obligatorily expressed as a syntactic argument. Speakers tend to produce clauses presenting

situations which single out one participant. The subject is the argument that is singled out as syntactically prominent, and hence is a good choice to express the referent to which the “aboutness condition” applies. Bresnan’s (2001:98) classification of subject as a grammaticalised discourse function is based on these factors; she also points out that c-structure properties can reflect this prominence, in that subjects as well as topics and foci can be required to precede or c-command other constituents in the clause.

The correlation between subjects and topics can be manifested in a number of ways, more strongly in some languages than in others. We have seen that topical marking in some languages is restricted to subjects. Other languages impose semantic constraints on subjects related to the notion of “topic-worthiness” (see Chapter 3, Section 3.2.3): only subjects with topic-worthy features are acceptable. For instance, Givón (1976) notes that in Malagasy and Kinyarwanda, subjects must be definite or generic. But even in languages without semantic restrictions on subjects, such as English, the overwhelming majority of subjects are definite. Textual counts demonstrate that in coherent discourse, the majority of subjects are pronominal, and unaccented pronouns are normally topical (Lambrecht 1987 and others).

Some languages place even stronger restrictions on the relation between syntax and information structure: they disallow nontopical subjects altogether. For example, Bresnan and Mchombo (1987:778), citing Bokamba (1981), show that in Dzamba (Bantu), all nonsubject arguments can be questioned in situ, but the subject cannot be questioned in its initial position: it must be questioned using a headed or headless relative clause.

- (19) a. \*Nzányi ó-wimol-aki ó-Biko e-kondo loo mé?  
 [Who told Biko a story/tale today?]
- b. ó-Moto ó-wimol-aki ó-Biko e-kondo loo mé nzányi?  
 The person who told Biko a story/tale today is who?  
 (Bresnan and Mchombo 1987:778)

Since wh-questions are inherently associated with narrow focus (see Chapter 3, Section 3.2.1), this indicates that Dzamba does not allow focused subjects, and indeed Bresnan and Mchombo (1987) claim that Dzamba subjects are grammaticalised topics and for this reason cannot bear the focus role.

Information structure role may affect the grammatical realisation of arguments. We will discuss the effect of topicality on the realisation of arguments as objects in Chapter 9; here we briefly discuss its effect on the realisation of subjects. In some languages, a nontopical or focused argument may not be realised as subject, but must be demoted to a lower function on the hierarchy. According to Lambrecht (2000), the functional motivation for this is “paradigmatic contrast”: topicless sentence-focus constructions take on a form which

distinguishes them from topic-comment structures in which the subject is topical.

In particular, the thematically highest argument of the verb may be assigned to different grammatical functions, depending on its information structure role. In constructions where it is not topical, the thematically highest argument tends to be associated with the behavioural features of objects. Creissels (2008) shows that in French impersonal constructions, the argument of an intransitive verb can appear postverbally:

- (20) a. *une femme viendra*  
         a woman come.Fut.3Sg  
         ‘A woman will come.’
- b. *il viendra une femme*  
         Expl come.Fut.3Sg a woman  
         ‘A woman will come (there will be a woman coming).’
- (Creissels 2008:155)

Here the postverbal NP is located in the canonical object position and patterns with objects, rather than subjects, with respect to a number of syntactic properties. Creissels concludes that it is a syntactic object, although it has the same semantic role as the canonical subject in (20a), and Cummins (2000) argues for the same conclusion. The split is motivated by information structure considerations: while the subject in (20a) is topical, (20b) is a presentational “all-new” construction, and the postverbal NP belongs to the broad focus domain. Creissels (2008:157) refers to such cases as “pragmatic conditioning on fluid transitivity”.<sup>4</sup>

Nikolaeva (2001) argues that in transitive clauses in Northern Ostyak (Uralic) the subject must bear the primary topic role. When the agent argument is non-topical, passivisation is required. This claim is supported by the following observations. First, topicalisation of a non-agent argument triggers passivisation. In passive clauses, the agent is encoded as a locative NP, while another topical argument is the subject. Kulonen (1989) shows that arguments with different semantic roles can be realised as the subject in a passive construction: patient/theme, recipient/benefactive, location, goal, and time. The examples in (21) demonstrate that the passive construction is required when the primary

<sup>4</sup>Following work by Platzack (1983) and others, Lødrup (1999) proposes a similar analysis for the presentational focus construction in Mainland Scandinavian, which involves an expletive in subject position and a postverbal, presentationally focused NP. However, Börjars and Vincent (2005) argue for an alternative analysis of this construction in which both the expletive and the presentationally focused postverbal NP bear the subject role.

topic, the element that bears the aboutness relationship to the proposition, is not associated with the agent argument.

- (21) a. What about Peter?  
 b. (luw) Juwan-na re:sk-əs-a  
 he John-Loc hit-Past-Pas.3SgSubj  
 ‘John hit him.’ (literally: ‘He was hit by John.’)  
 c. (luw) Juwan-na ke:si-na ma-s-a  
 he John-Loc knife-Loc give-Past-Pas.3SgSubj  
 ‘John gave him a knife.’ (literally: ‘He was given a knife by John.’)

Unlike English, the use of the non-passive construction in this context is strictly ungrammatical; the active counterparts of (21b,c) cannot be used in this context:

- (22) a. \*Juwan (luwe:l) re:skə-s-li  
 John he.Acc hit-Past-Obj.3Sg.Subj  
 ‘John hit him.’  
 b. \*Juwan luw e:lti ke:si ma-s  
 John he to knife give-Past.3Sg.Subj  
 ‘John gave him a knife.’

Additional evidence for the topichood of the subject is provided by sentences with a focused agent. Questions and answers involving the agent require passivisation.

- (23) a. kalaŋ xoj-na we:l-s-a?  
 reindeer who-Loc kill-Past-Pas.3SgSubj  
 ‘Who killed the reindeer?’  
 b. Juwan-na we:l-s-a  
 John-Loc kill-Past-Pas.3SgSubj  
 ‘John did.’

The ungrammaticality of the non-passive counterparts of (23) in this context immediately follows from the requirement for the subject to be associated with topic, together with the assumption that focus and topic cannot correspond to the same sentence element. Finally, as mentioned above, topic expressions

must be referential. In Ostyak, quantified expressions such as *anybody* or *nobody* do not occur as subjects of transitive clauses. As example (24) demonstrates, when they correspond to an agent-like argument, the clause must be passivised.

- (24) a. tam xu:j xoj-na an wa:n-s-a  
 this man who-Loc Neg see-Past-Pas.3SgSubj  
 ‘Nobody saw this man.’
- b. \*xoj tam xu:j an wa:nt-əs / wa:nt-əs-li  
 who this man Neg see-Past.3SgSubj see-Past-Obj.3SgSubj  
 ‘Nobody saw this man.’

These examples show that Ostyak grammar displays a strong correlation between the grammatical function of subject and the (primary) topic role.

Other languages are less strict and allow for nontopical subjects, but only when special devices are used. That is, the subject referent is interpreted as topic unless there are syntactic or prosodic cues to the contrary, such as appearance of a topical nonsubject element in a noncanonical sentence-initial position, or prosodic prominence of focus subjects in event-reporting (sentence focus) sentences. The latter is illustrated by the following contrast from Lambrecht (1994).

- (25) a. How’s your neck?  
 My neck/It HURTS.
- b. What’s the matter?  
 My NECK hurts. (Lambrecht 1994:137)

The answer in (25a) is a predicate focus structure: the comment *hurts* provides new information about the topical referent under discussion (neck). In contrast, (25b) is an event-reporting sentence focus (orthetic) utterance: here, the new information is associated with the whole proposition, so the subject is not topical. This difference is expressed formally: the topical subject corresponds to an unaccented NP and can be pronominalised, while the nontopical subject is a prosodically prominent lexical NP. The subject in (25b), as opposed to (25a), is structurally and functionally marked.

#### 5.4 Conclusion

We have shown that primary topicality can be expressed by case or agreement, and that in some languages this marking is compatible with any of several grammatical functions. However, unmarked primary topics are subjects, and

in many languages the marking of primary topicality is restricted to subjects: topical subjects are associated with more grammatical marking (case and/or agreement) than nontopical subjects. We also find languages where nontopical arguments cannot appear as subjects, but must be demoted through voice-like alternations.

## 6

# *Topical marking of nonsubjects*

Though the correlation between topicality and grammatical marking has been well studied for subjects, it is less well studied for nonsubject topics. The languages we discuss in this chapter show clearly that topicality can be a relevant factor in determining patterns of grammatical marking of nonsubjects. In these languages, marking is associated with sentence elements bearing any one of a variety of grammatical functions other than subject, but is determined partially or completely by reference to topichood. In this sense, these patterns are similar to those discussed in Chapter 5, Section 5.1, except that in the languages discussed here, marking is permitted only for nonsubjects, and disallowed for subjects.

### **6.1 Casemarking of topical nonsubjects**

#### *6.1.1 Persian*

The Persian (Iranian) postposition *râ* (with various colloquial and dialectal variants) marks arguments bearing various grammatical functions, and its distribution has often been argued to be conditioned by information-structural factors. First, certain time adverbials can be marked by *râ*, such as, for example, *hafteye âyanda* ‘next week’, *emšæb* ‘tonight’ or *tabestan* ‘in summer’.

- (1) a. *kæmal emšæb-o inja mi-mun-e*  
 Kamal tonight-RA here Impf-stay-3Sg  
 ‘Tonight Kamal is staying here.’ (Mahootian 1997:121)
- b. *hafte-ye âyanda-ro esterâhat mi-kon-am*  
 week-Ez coming-RA relax Impf-do-1Sg  
 ‘As for the next week, I will relax.’ (Karimi 1990:143)

Furthermore, *râ* may be present on floating topics located at the left periphery of the sentence and cross-referenced by a clitic. The floating topic can correspond to a number of different grammatical functions. In (2a), it corresponds to an oblique object of the verb ‘laugh’, while in (2b) it corresponds to a comitative adjunct.

- (2) a. *man-o beh-me mi-xand-e*  
 I-RA at-1Sg Impf-laugh-3Sg  
 ‘She laughed at me.’ (Karimi 1990:143)
- b. *sâsân-o bâh-a raqsid-am*  
 Sasan-RA with-3Sg dance.Past-1Sg  
 ‘Sasan, I danced with him.’ (Karimi 1990:154)

The floating topic phrase is often associated with the possessor of the object, which is also marked by *râ*.

- (3) *mašin-o dar-eš-o bast-am*  
 car-RA door-3Sg-RA close.Past-1Sg  
 ‘As for the car, I closed its door.’ (Karimi 1990:143)

Finally, *râ* is a frequent object marker:

- (4) *man ketâb-râ xarid-am*  
 I book-RA buy.Past-1Sg  
 ‘I bought the book.’ (Dabir-Moghaddam 1992:557)

Thus, the postposition *râ* marks direct objects, adverbials and floating topics. In all of these functions, *râ* is optional. We maintain that its distribution on non-objects is determined at least in part by information structure: *râ* appears on nonsubject topics as well as on some nonsubject phrases with features of topic-worthiness, such as definiteness.<sup>1</sup>

<sup>1</sup>Dabir-Moghaddam (1992) mentions one “exceptional” case where *râ* appears on the main clause subject and also corresponds to a nonsubject gap within the relative clause. We do not know how this example fits with the overall distribution of *râ*.



Adverbials marked with *râ* must be interpreted as topical. The relevant utterances are construed as information about the relation between the subject referent and a certain time span which delimits the action performed by the subject referent. For instance, in example (5) summer is mentioned in the previous context, and so the speaker has reason to believe that the addressee considers it important and can expect a certain pragmatic relation between summer and the subject referent to be expressed. The unmarked adverbial *tabestan* is impossible or strongly dispreferred:<sup>2</sup>

- (5) a. What are your plans for the summer?  
 b. *tabestan-râ/?\*tabestan esterâhat mi-kon-am*  
*in.summer-RA/ in.summer relax Impf-do-1Sg*  
 ‘In summer/as for the summer, I will relax.’

Example (5) contrasts with example (6):

- (6) a. When will you finally relax?  
 b. *tabestan/?\*tabestan-râ esterâhat mi-kon-am*  
*in.summer/ in.summer-RA relax Impf-do-1Sg*  
 ‘I will relax in summer.’

In example (6), summer is not pragmatically presupposed and constitutes part of the new information associated with the utterance, so the adverbial ‘in summer’ is in focus; here, the distribution of *râ* is opposite to (5).

Similar considerations apply to floating topics. When a constituent marked by *râ* is a floating topic, as in (2) and (3), fronting (topicalisation) is an additional clue to its topical status. Oblique objects or possessors are infrequent topics, but can be interpreted as such under certain pragmatic conditions. Topical possessors marked with *râ* tend to be inalienable; this is because a statement about an inalienably possessed entity also contributes information about its possessor, without which the inalienably possessed object cannot be conceptualised. According to our consultants, (3) is possible only in a context where the relation between the subject referent and the car (possessor of the object ‘door’) is under discussion. For example, it can be construed as an answer to the question ‘What did you do to the car?’ (‘What I did to the car is, I closed its door’), but is an inappropriate answer to ‘What have you just done?’, which does not presuppose any pragmatic association between the speaker and the car. The referent of the object, ‘door’, is less pragmatically important than the referent of the external possessor ‘car’. The utterance is not construed as

<sup>2</sup>Examples for which the source is not mentioned come from personal communication with Ghazaleh Khad and Shamsi Saber.

being about the relation between the subject referent and the object ‘door’, but rather about the relation between the subject referent and the possessor ‘car’. Therefore it cannot answer the question ‘What did you do to the door of the car?’.

The situation with direct objects is more complicated, and various theories have been proposed for the distribution of *râ* on objects. Windfuhr (1979) was among the first to suggest that *râ* marking has to do with topicality, while Browne (1970) and Karimi (1990) claim that specificity is the relevant factor. Lazard (1992) argues for the importance of the degree of individuation of the *râ*-marked argument: the more the object referent is individuated and affected by the action designated by the verb, the more likely it is that *râ* is used. According to Ghomeshi (1997) and Bossong (1991:64), several properties are relevant for *râ* marking. These include topicality as well as animacy, definiteness, and affectedness.

We are in agreement with the view that the distribution of *râ* on objects is at least partly explainable in terms of topicality. In this we follow Dabir-Moghaddam (1992), who proposed that the main function of *râ* is what he calls secondary topic marking. Though he does not provide a precise definition of secondary topic, his understanding of this notion seems to be close to ours, judging from the contexts he provides for his examples. We note two important differences between his analysis and ours. First, the examples in (7) indicate that *râ* can mark the primary topic as well as the secondary topic. The subjects in (7) are focused, and since there are only two NPs in each of the examples, the *râ*-marked object cannot be the secondary topic; rather, it must be the primary topic:

- (7) a. ki mašin-i-ra/?\*mašin-i did  
 who car-Indef-RA/car-Indef see.Past.3Sg  
 ‘Who saw a car?’
- b. ki ketâb-i-ra/?\*ketab-i xarid  
 who book-Indef-RA/book-Indef buy.Past.3Sg  
 ‘Who bought a book?’

Second, we disagree with Dabir-Moghaddam’s position that topicality is the only relevant factor in determining *râ* marking: we believe that topicality is a factor for some objects, while on other objects the motivation is essentially semantic, having to do with features of topic-worthiness.

In particular, *râ*-marking works differently for definite and indefinite objects. As shown by Lazard (1982), all definite objects must be marked, independent of their information structure function. The omission of *râ* in (4) with the given translation yields strict ungrammaticality:

- (8) a. What are you eating?  
 b. in sib-o/\*sib mi-xor-am  
 this apple-RA/ apple Impf-eat-1Sg  
 ‘I am eating this apple.’ (Karimi 1990:148)

For the same reason, all personal pronouns, proper nouns and objects with demonstratives must be marked even if they are in focus (Ghomeshi 1997:137). Thus, information structure motivations are irrelevant for definite objects; they must be marked due to their definite status.

This observation accounts for the fact that *râ* can appear twice in a single clause. As evidenced by example (2b), *râ* can be found on both an object and its possessor. In (9), *râ* marks the definite object as well as the temporal expression ‘just this one hour’.

- (9) faqat in ye sâ’at-o in ketâb-o be-xun  
 just this one hour-RA this book-RA Imp-read  
 ‘Read this book this one hour!’ (Ghomeshi 1997:151)

Karimi (1990) takes examples like (9) to contradict the topic analysis of *râ*, on the basis of her assumption that the topic role is unique in a clause. As we have seen, however, topical arguments need not be unique in the clause: Chapter 3 introduced a distinction between the primary topic and the secondary topic, and we assume that the presence of a secondary topic entails the presence of a primary topic. Furthermore, in constructions like (9) the object is definite, and definite objects must be marked independently of any informational requirements. In fact, the first and the second *râ* in (9) have different functions: the first marks the topicality of the temporal adjunct, while the second is licensed by definiteness.

For indefinite objects, *râ* is optional. According to Lazard (1982), indefinites that require *râ* are either partitive or have a “a certain X” interpretation. That is, they must be “pragmatically anchored” in the sense of Lambrecht (1994) or “specific” in the sense of Enç (1986) (see Chapter 3, Section 3.2.2): the referent is chosen from a familiar set and linked to entities in the domain of previous discourse. Indeed, nonspecific (nonreferential) objects cannot bear *râ*. But the correlation between *râ*-marking on indefinite objects and specificity is imperfect: although indefinite objects must be specific to be marked by *râ*, not all specific objects are *râ*-marked.

This can be shown by testing for specificity in an intensional context, using anaphoric pronouns like *it* or *one*. If the antecedent is specific, the anaphor must be definite, whereas nonspecific antecedents can be referred to by the indefinite anaphor ‘one’. As Ghomeshi (1997:138–139) shows, this test indicates that indefinite objects without *râ* can be specific.

- (10) *jiân mi-xâst ye qalam-i peyda kon-e va peyda-š*  
 John Impf-want.3Sg one pen-Indef find do-3Sg and find-3Sg  
*kard*  
 do.Past.3Sg  
 ‘John wanted to find a pen and he found it.’ (Ghameshi 1997:139)

In (10), the pronominal object clitic *-š* ‘it’ in the second conjunct refers to the pen John wanted to find, and so the NP *ye qalami* ‘a pen’ must be construed as specific. Nevertheless, it does not host *râ*.

Recall Erteschik-Shir’s (2007) examples from Danish involving indefinite objects with and without relative clauses, cited in Chapter 3, Section 3.2.2 (example 7, page 54); we argued there that the relative clause renders the indefinite object specific and therefore potentially topical. The situation in Persian seems to be analogous: *râ* marks a specific indefinite object as topical, and if an indefinite object is not perceived as pragmatically salient, *râ*-marking is impossible. Consider the following contrast:

- (11) a. *?/\*man sib-i-râ xord-am*  
 I apple-Indef-RA eat.Past-1Sg  
 ‘I ate an apple.’  
 b. *man sib-i-râ xord-am ke az deraxt oftadebud*  
 I apple-Indef-RA eat.Past-1Sg that from tree fell.Past.3Sg  
 ‘I ate an apple that fell from the tree.’

In (11a), object marking on an indefinite object is impossible. In (11b) it is possible, though not required. This is because the object is interpreted as specific and therefore can (although need not) be topical.

In the examples discussed so far, the relation of a specific indefinite to another familiar entity is expressed in the same sentence. In example (11b) it is expressed by a relative clause, while in example (6) of Chapter 3 it is expressed by a modificational PP. However, in some cases this relationship seems to be pragmatically assigned by the speaker based on extrasentential factors. In particular, an indefinite object may be interpreted as topical if it is mentioned in the context immediately following the relevant utterance. Dabir-Moghaddam shows that if *râ* is hosted by an indefinite object, the subsequent discourse must provide more information about the object referent, for example by means of a coordinate structure.

- (12) a. *man ketâb-i/\*ketâb-i-râ xarid-am*  
 I book-Indef/ book-Indef-RA buy.Past-1Sg  
 ‘I bought a book.’

- b. man ketâb-i/ketâb-i-râ                    xarid-am    va ...  
 I    book-Indef/ book-Indef-RA buy.Past-1Sg and  
 ‘I bought a book and ...’                    (Dabir-Moghaddam 1992:557)

An appropriate continuation for the sentence in (12) is something like *sent it to my brother*. Without this elaboration, a sentence with a *râ*-marked indefinite object is intuitively incomplete, and speakers perceive it as unacceptable. Although in (12b) the object referent is new to the addressee and marked as indefinite, it receives grammatical marking due to its topical status. The topical status of the object is licensed by the fact that its referent plays a role in the following discourse. We will show in Chapter 8, Section 8.3.1 that this condition is not unique to Persian: topicality marking on objects in Khalkha Mongolian works in very much the same way.

In sum, although marking on direct objects correlates with topicality, this correlation is not absolute, unlike for adverbials and floating topics. The postposition *râ* marks all definite objects as well as topical indefinite objects. In Chapter 10, we show that this situation is typical of other languages in which DOM is conditioned partly by information structure and partly by semantics, and we suggest that this situation can result from historical spreading of *râ*-marking to nontopical definite objects with features typical of topics.

### 6.1.2 Tariana

Tariana, an Arawakan language described by Aikhenvald (2003), indicates the topicality of nonsubject elements by means of the suffix *-nuku/-naku*. Aikhenvald is not completely consistent in identifying the status of this marker: it is sometimes referred to as a case marker indicating a syntactic function (Aikhenvald 2003:158) and sometimes as a “topical non-subject marker” (Aikhenvald 2003:160). It can appear on objects, which are not otherwise casemarked, or on non-objects together with casemarking indicating syntactic function; we will treat it as indicating topicality alone. It can also appear on dependent verbs, in which case it indicates that the subordinate clause headed by the dependent verb contributes background information: we will not discuss these cases here.

According to Aikhenvald, objects marked with *-nuku/-naku* must satisfy at least one of the following conditions: (i) they must be the topic of a narrative, (ii) they must be referential, specific and/or definite, and (iii) they must be pragmatically important. Parameter (i) is a strong indication of topicality in our sense, on the assumption that discourse topics are normally interpreted as sentence topics (though not the other way around). Parameter (iii) also strongly suggests topicality. It remains unclear how (ii) interacts with the other two conditions — in particular, whether each of the relevant properties is sufficient for topical marking of nonsubjects, or whether the combination of properties

in (ii) reflects a minimal condition that all marked nonsubjects must share. No discussion is provided of cases where the properties listed in (ii) conflict with each other: for example, there are no examples with specific indefinite objects. Still, the general tendency is clear: objects marked with *-nuku/-naku* are topical. Example (13) is part of a story about gold miners; here, ‘gold’ is topical and marked with *nuku*:

- (13) di-hē-ta-pidaha diha  
 3Sg.Non.Fem-see.Caus1-Caus2-Remote.Past.Reported Art  
 paiku-nuku  
 gold-Top  
 ‘He showed the gold.’ (Aikhenvald 2003:145)

Nontopical objects are unmarked if lexical, or marked with a different marker, *-na*, if pronominal. The object ‘women’ in example (14) is nontopical and unmarked:

- (14) mhaīda ina pi-awada  
 Proh woman.PI 2Sg-think.about  
 ‘Don’t think about women.’ (Aikhenvald 2003:153)

In example (15), ‘payment’ is nontopical and unmarked, while ‘secrets’ is topical and marked:

- (15) di-weni na-na-ka-pidana  
 3Sg.Non.Fem-pay 3PI-want-Sub-Remote.Past.Reported  
 na-pia-nipe-nuku na-kalite-pidana du-na  
 3PI-hide-Nmlz-Top 3PI-tell-Remote.Past.Reported 3Sg.Fem-Obj  
 ‘Wanting a payment, they told their secrets.’ (Aikhenvald 2003:145)

Importantly, *-nuku/-naku* marking is not restricted to objects, but occurs on other grammatical functions if they are topical. For instance, it may appear on topical instrumental and locative obliques. Aikhenvald analyses such cases as involving “double case marking”, although these cases are clearly different from canonical instances of “case stacking” (Plank 1995, Nordlinger 1998): one of the two “case markers” is actually a marker of information structure function, similar to the Japanese examples with grammatical marking followed by topic marking discussed in Chapter 5, Section 5.1.

- (16) a. Then she took the pestle, she carefully took her son, and pulled him out (of a hammock).
- b. diha-da        ye:da-ne-nuku dhuepaneta  
 Art-Cl:Round pestle-Inst-Top 3SgFem.exchange  
 du-kwe-ta-pidana  
 3SgFem-hand-Caus-Caus-Remote.Past.Rep  
 ‘She exchanged (the child) with the pestle and hung it.’  
 (Aikhenvald 2003:159)

The context in (16a) establishes the salient role of the referent ‘pestle’ in (16b), an instrumental oblique which is also marked for topicality. An external topic which is not an argument of the clause can also be marked with *-nuku/-naku*:

- (17) nhua-naku kida-mhana  
 I-Top        ready-RemotePast.NonVisual  
 ‘For me, it was over.’ (Aikhenvald 2003:146)

Like Persian *râ*, *-nuku/-naku* marking is never found on subjects. In fact, subject topicality is not formally marked in Tariana, though a subject may receive overt marking for focus. Example (18) demonstrates that *-nuku* marking on the object may be present if the subject is in focus (and marked as such):

- (18) wha-ne wa-kesi-pe-naku hipay yapise-se-naku  
 we-Foc 1Pl-relative-Pl-Top ground under-Loc-Top  
 ma-pe-kade-naka  
 Neg-throw-Neg-Pres.Visual  
 ‘We don’t want to throw our relatives under the ground.’ (Aikhenvald 2003:636)

This example also demonstrates that *nuku/-naku* can mark primary as well as secondary topics: in example (18) the subject has focus marking and cannot be the primary topic, and Aikhenvald notes that in the context in which this sentence was produced, ‘relatives’ is topical.

Aikhenvald also presents examples showing that *-nuku/-naku* marking can appear twice in the same clause. In (19), it marks the object as well as the (inalienable) possessor of the object:

- (19) diha-pasi-nuku di-whida-nuku        du-pisa-taka  
 he-Aug-Top    3Sg.Non.Fem-head-Top 3Sg.Fem-cut-off  
 du-pe  
 3Sg.Fem-leave  
 ‘She cut off the head of him, the big one.’ (Aikhenvald 2003:157)

In (19), the possessor and the possessed need not be adjacent. The possessor does not have object properties, and does not seem to have been promoted to a role within the main clause: for example, it cannot be passivised. It may be that the first *nuku*-marked phrase in this construction is a floating, clause-initial topic that is interpreted as the possessor of the object, similar to the Persian examples in (2) and (3) above.

### 6.1.3 *Topical nonsubject casemarking*

The Tariana nonsubject topic marker *-nuku/-naku* is associated with almost the same specifications as for the Quechua and Japanese markers discussed in Chapter 5, except that the marking may appear on any argument except the subject:

(20) Topic marking (nonsubjects), Tariana:

$$\begin{aligned} &\neg(\text{SUBJ } \uparrow) \\ &(\uparrow_{\sigma} \text{ DF}) = \text{TOPIC} \end{aligned}$$

The first of these two specifications is similar to the requirement associated with topical subject marking, presented in (17) of Chapter 5, except for the crucial presence of the negation operator. The requirement  $\neg(\text{SUBJ } \uparrow)$  ensures that the argument bearing topical casemarking (the f-structure  $\uparrow$ ) is **not** the subject of its clause; it may bear any other grammatical function. As in (17) of Chapter 5, the second line specifies that the casemarked argument's discourse function is TOPIC, as described in Chapter 4, Section 4.3.

The specifications associated with Persian *râ* are more complex, since *râ* is not invariably associated with topics; it marks definite objects whether or not they are topical, as well as arguments that are demonstrably topical: indefinite objects, some adjuncts, and external/displaced topics, which we treat as bearing the f-structure role of grammaticalised TOPIC (for discussion of the grammaticalised TOPIC role at f-structure for displaced constituents, see Chapter 4, Section 4.1.2). These constraints are captured in the following specifications:

(21) Topic marking (nonsubjects), Persian:

$$\begin{aligned} &\{(\text{OBJ } \uparrow) \\ &(\uparrow \text{ DEF}) = + \\ &|(\{\text{TOPIC|OBJ|ADJ} \in\} \uparrow) \\ &(\uparrow_{\sigma} \text{ DF}) = \text{TOPIC}\} \end{aligned}$$

This specification is disjunctive, reflecting the dual nature of *râ* as a definite object marker and a marker of topical elements. The disjunction is surrounded by curly brackets, and the two parts are separated by a vertical stroke. The



first part consists of the first two lines, which specify that *râ* can mark definite objects: the first line requires the argument marked with *râ* to be an object, and the second line requires it to be definite (with the feature DEF +). There is no requirement for *râ*-marked definite objects to be topical; in Chapter 10, we discuss the historical relation between topicality and semantic features of topic-worthiness, including definiteness. The second part, the third and fourth lines in (21), requires the *râ*-marked argument to be either a displaced f-structure TOPIC, an object, or an oblique, and to bear the information structure role of TOPIC, as in the Tariana specifications in (20).

## 6.2 Agreement with topical nonsubjects

### 6.2.1 *Itelmen*

In *Itelmen* (Chukotko-Kamchatkan) as described by Bobaljik and Wurmbrand (2002), subject and direct object arguments are not casemarked and can be omitted under pro-drop. The *Itelmen* verb has two agreement slots, primary (prefixal) agreement and secondary (suffixal) agreement.<sup>3</sup> The basic structure of the verbal form is as follows:

(22) Agreement1-Mood-Stem-Aspect-Tense-Agreement2

Prefixal agreement (Agreement1) references the subject and can sometimes be null. We are concerned here with suffixal agreement (Agreement2), which is influenced by information structural factors: when there is a choice of agreement controller, the *Itelmen* verb agrees with a nonsubject topic.

Suffixal agreement in *Itelmen* is obligatory; following Bobaljik and Wurmbrand, we take this to be a morphological fact. There are two distinct sets of agreement suffixes for third person nonsubject elements: portmanteau suffixes express the features of a third person direct object and, under some circumstances, the subject, while a separate set of suffixes simply reference a third person oblique. Suffixal agreement with first and second person nonsubjects never makes reference to the subject. We will gloss the agreement morphemes as either *Obj* or *Obl*, depending on the grammatical function of the agreement controller.

Bobaljik and Wurmbrand show that exponence of the features of the subject in suffixal agreement arises in two situations. In intransitive clauses, subject agreement can be expressed twice, once in the prefixal slot and again in the suffixal slot:

---

<sup>3</sup>This oversimplifies the morphology of the *Itelmen* verb; for more detailed discussion of *Itelmen* verb morphology and agreement, see Bobaljik and Wurmbrand (2002).

- (23) kma t-k'ot-kičen  
 I 1SgSubj-come-1SgSubj  
 'I came/arrived.' (Bobaljik and Wurmbrand 2002:(6a))

When the direct object is third person, object suffixal agreement can express features of both the direct object and the subject:

- (24) kza əlčqu-n na  
 you see-2SgSubj>3SgObj him  
 'You saw him.' (Bobaljik and Wurmbrand 2002:(12d))

It is also possible for the verb to show oblique suffixal agreement, which is in complementary distribution with the object agreement suffixes, and does not involve expression of subject features. The following examples illustrate secondary oblique agreement: example (25a) shows agreement with an oblique source argument, which appears with dative/locative casemarking, while (25b) shows secondary oblique agreement with the otherwise unexpressed possessor of the subject.

- (25) a. kma iplɣ-enk t-ɬ-nen βəpɣ-eʔn  
 I friend-Dat/Loc 1SgSubj-take-3SgObl fly.agaric-Pl  
 'I took fly agaric (mushrooms) from my friend.'  
 (Bobaljik and Wurmbrand 2002:(14c))
- b. da quskɫnaqu ɬʲɣ<sup>w</sup>i ploxo le-ɣ<sup>w</sup>in ktxiŋ qaʔt  
 Interj Q. very bad become-3SgSubj head already  
 iʔtɬe-s-kinen  
 split-Pres-3SgObl  
 'And Q. began to feel very bad, already his head is splitting.'  
 (Bobaljik and Wurmbrand 2002:(32b))

If there are several potential nonsubject agreement controllers in a clause, the choice of agreement controller is determined by pragmatic factors. In (26), the verb 'give' agrees with the subject and one of its two nonsubject arguments, with the object competing with the oblique for the agreement slot:

- (26) a. isx-enk n-zəl-aɬ-um kza kəma-nk  
 father-Loc Imprp-give-Fut-1SgObj/Obl you me-Dat  
 'Will father give you to me?' (Bobaljik and Wurmbrand 2002:(14b))
- b. isx-enk n-zəl-aɬ-in kza kəma-nk  
 father-Loc Imprp-give-Fut-2SgObj/Obl you me-Dat  
 'Will father give you to me?' (Bobaljik and Wurmbrand 2002:(15))

In (26a) secondary agreement is with the dative indirect object ‘me’, and in (26b) agreement is with the direct object ‘you’. The translation is roughly the same, and Bobaljik and Wurmbrand (2002) emphasise that there is no evidence from casemarking, word order, or other syntactic criteria that these different agreement patterns are evidence of an alternation in grammatical functions. Instead, they demonstrate convincingly that conditions on the choice of the secondary agreement controller are “governed by discourse-pragmatic considerations such as salience”. We believe that this indicates that secondary agreement depends on information structure: when there is a choice of agreement controllers, the verb agrees with topical nonsubjects.

The influence of topicality is clear when contextual information is provided. Bobaljik and Wurmbrand provide example (27), in which the direct object/theme, the controller of verb agreement, is established as topical in the context provided by the preceding sentence:

- (27) maʔ kəman βaʔčʔ k'e-nk t-zəl-čen?  
 where my knife who-Dat/Loc 1SgSubj-give-1SgSubj>3SubjObj  
 ‘Where is my knife? Who did I give it to?’  
 (Bobaljik and Wurmbrand 2002:(18b))

In this example the knife is under discussion when the second clause is produced, so the object NP triggers secondary agreement. The same verb agrees with the indirect object in example (28), where the previous context establishes the goal rather than the theme as topical:

- (28) zlatumx piki-in. əŋqa ənna-nk t-zəl-nen?  
 brother go-3SgSubj what him-Dat/Loc 1SgSubj-give-3SgObl  
 ‘My brother left. What did I give to him?’  
 (Bobaljik and Wurmbrand 2002:(19b))

The same distribution is observed in sentences without wh-question words.

- (29) a. maʔ βaʔčʔ qeʔnu zlatumx-enk  
 where knife really brother-Dat/Loc  
 t-zəl-čen  
 1SgSubj-give-1SgSubj>3SubjObj  
 ‘Where is the knife? Didn’t I give it to my brother?’  
 (Bobaljik and Wurmbrand 2002:(20a))
- b. i kma ənna-nk βaʔč t-zəl-nen  
 and I him-Dat/Loc knife 1SgSubj-give-3SgObl  
 [My brother came]. ‘And I gave the knife to him.’  
 (Bobaljik and Wurmbrand 2002:(20b))

The context makes it clear that the topic role is associated with the direct object in (29a) and with the dative oblique in (29b).

Example (25b) above, where agreement is with the possessor/dative argument, additionally shows that the secondary agreement controller is topical. The discourse fragment consists of two clauses, and the character called Qusklnaqu is the topic of the first clause. The second clause states that Qusklnaqu's head is splitting. Since the head is inalienably possessed by Qusklnaqu, the second clause provides new information about him. In this context, both referents (Qusklnaqu and his head) are under discussion at the time of the utterance. Note that the presence or absence of overt (pronominal) agreement controllers does not affect agreement. Thus, it is information structure rather than grammatical role that determines nonsubject third person agreement in Itelmen: the verb agrees with nonsubject elements that are topical.

The influence of the person hierarchy on agreement in Itelmen is not yet clear. Bobaljik and Wurmbrand point out that in the related languages Alutor and Chukchi, agreement is required with first or second person agreement controllers. However, they demonstrate that in Itelmen, this is only a strong tendency and not a firm requirement; they provide example (30), in which suffixal agreement is controlled by a third person element even in the presence of a potential second person controller:

- (30) kma xe<sup>l</sup>?nč kn-ank nənč m-zəl-čən  
 I not you-Dat fish 1SgSubj-give-1>3SgObj  
 'I won't give the fish to you.' (Bobaljik and Wurmbrand 2002:(28))

Pending further research, we leave open the question of the influence of person on the determination of agreement controller in Itelmen.

### 6.2.2 *Tabassaran*

Kibrik and Seleznev (1980) discuss two types of person agreement in Tabassaran (North Caucasian): the first type is subject agreement, and the second type is with a nonsubject element which is pragmatically prominent. In the first type of agreement, the verb agrees with a first or second person subject which can bear various semantic roles and can be casemarked in several ways (normally Nominative, Dative or Ergative). With a third person subject, the verb may receive the affix *-(u)v*. Kibrik and Seleznev analyse this affix as a default non-agreement marker rather than third person agreement, but we find their arguments unclear, and for simplicity we will treat this affix as expressing third person agreement.

Any prominent nonsubject argument can trigger the second type of agreement, including patient and recipient as well as (arguably) some non-terms.

Nonsubject agreement is realised differently depending on the case of the agreement controller: for example, agreement with a first person singular argument is marked either with the affix *-is* or with *-za/-zu*, depending on whether the argument stands in the dative or nominative/ergative, respectively. There are additional morphological constraints on the realisation of agreement features in nonsubject agreement, the details of which are not relevant for present purposes.

In some circumstances, the verb can agree with a subject as well as a prominent nonsubject argument, combining the first and second types of agreement. In particular, the verb can take two agreement affixes when the subject bears the agent role and is in the first person; in such circumstances, the verb obligatorily hosts subject agreement, and optionally hosts nonsubject agreement with a prominent nonsubject element, as in example (31):

- (31) izu ivu uvčunu-zu-vu  
 I.Erg you.Nom beat-1Sg.Nom/Erg-2Sg.Nom/Erg  
 ‘I am beating you.’ (Kibrik and Seleznev 1980:20)

In all other cases only one agreement affix is present, either subject agreement or nonsubject agreement. The examples in (32) have third person subjects, and therefore only one agreement affix is allowed: either the third person agreement suffix *-(u)v*, or a suffix encoding agreement with a pragmatically prominent nonsubject element.

- (32) a. duRu izu uvčun-uv / uvčunu-za  
 he.Erg I.Nom beat-3 beat-1Sg.Nom  
 ‘He has beaten me.’ (Kibrik and Seleznev 1980:23)
- b. duRu izus at’nar uRn-uv / uRn-is  
 she.Erg I.Dat socks.Nom knit-3 knit-1Sg.Dat  
 ‘She knitted socks for me.’ (Kibrik and Seleznev 1980:23)
- c. duRu iziʔin alarxun-uv / alarxunu-ziʔin  
 he.Erg I.Supess attack-3 attack-1Sg.Supess  
 ‘He attacked me.’ (Kibrik and Seleznev 1980:23)
- d. duRu izuq<sup>h</sup> hit’ik’in-uv / hit’ik’inu-zuq<sup>h</sup>  
 he.Erg I.Postess hide-3 hide-1Sg.Postess  
 ‘He hid behind me.’ (Kibrik and Seleznev 1980:23)

In (32a) nonsubject agreement is optionally triggered by the nominative patient argument, in (32b) by the dative recipient, in (32c) by an argument in the superessive (Supess) case, and in (32d) by the postessive (Postess) NP.

It could be argued that all of the examples in (32) involve predicate-argument agreement and therefore do not violate the usual assumptions about agreement domains. However, nonsubject agreement is also possible when the controller corresponds to a possessor, as shown in (33).

- (33) a. *jas agaji dumu uvčun-uv / uvčun-as*  
 I.Gen father.Erg he.Nom beat-3 beat-1Sg.Gen  
 ‘My father has beaten him.’ (Kibrik and Seleznev 1980:23)
- b. *duRu jas agaji-s k’až ik’v-uv / ik’n-as*  
 he.Erg I-Gen father-Dat letter.Nom write-3 write-1Sg.Gen  
 ‘He wrote a letter to my father.’ (Kibrik and Seleznev 1980:23)
- c. *baj jas č<sup>h</sup>huka-q<sup>h</sup> hit’ik’n-uv / hit’ik’n-as*  
 boy I.Gen shed-Postess hide-3 hide-1Sg.Gen  
 ‘The boy hid behind my shed.’ (Kibrik and Seleznev 1980:23)
- d. *duq’ari jas jak’u-xi hit’urd-uv / hit’urd-as*  
 they.Erg I.Gen axe-Com cut-3 cut-1Sg.Gen  
 ‘They are cutting (wood) with my axe.’  
 (Kibrik and Seleznev 1980:24)

In (33a) agreement is controlled by the possessor of the ergative subject, in (33b) by the possessor of the dative recipient, in (33c) by the possessor of the possessive NP, and in (33d) by the possessor of the comitative NP. In all these examples, nonsubject agreement is indicated by the first person singular genitive agreement affix *-as*.

Unfortunately, the precise conditions on nonsubject agreement are not clearly defined in the sources available to us. Kibrik and Seleznev (1980) state only that the agreement controller in nonsubject agreement is more “prominent” or “emphatic” than non-agreeing elements. For instance, if the verb in (32–33) shows agreement with the first person singular element, its pragmatic prominence is said to be assessed more highly than in corresponding sentences with (default) third person agreement. Based on their discussion, we cannot claim with certainty that agreement indicates topicality. Pragmatic prominence can be taken to indicate either that the agreeing element is salient in the sense of being under discussion (and thus topical), or that it is emphasised in the sense of being selected from a set of alternative candidates (and thus contrastive); we cannot resolve this issue without additional data, though we believe that it is cross-linguistically more common for agreement to encode topicality than

emphasis. For present purposes, the important point is that nonsubject agreement in Tabassaran is not restricted to a particular grammatical function and is conditioned by information structure, possibly topicality.<sup>4</sup>

### 6.2.3 Topical nonsubject agreement

Agreement marking involving nonsubject topics has the following general form, where GF is any grammatical function, and [GF–SUBJ] is any grammatical function other than subject:<sup>5</sup>

(34) General form of agreement with topical nonsubjects:

$$((\uparrow [\text{GF} - \text{SUBJ}])_{\sigma} \text{DF}) = \text{TOPIC}$$

This is also similar to the constraint for topical subject agreement, given in (18) of Chapter 5, except for the specification of the agreement controller: here, agreement is with any grammatical function **except** subject. We do not use this very general constraint for Itelmen and Tabassaran, however, since the agreement affix in these languages varies according to the case or grammatical function of the topical controller of agreement.

In Tabassaran, the agreement affix varies with the case of the agreement controller. Assuming that agreement does indeed mark topicality in Tabassaran, specifications for the first person singular dative agreement affix are as given in (35):

(35) Agreement with first person singular dative topical nonsubjects (Tabassaran):

$$\begin{aligned} (\uparrow [[\text{GF} (\text{POSS})] - \text{SUBJ}]) &= \%t \\ (\%t \text{ PERS}) &= 1 \\ (\%t \text{ NUM}) &= \text{SG} \\ (\%t \text{ CASE}) &= \text{DAT} \\ (\%t_{\sigma} \text{DF}) &= \text{TOPIC} \end{aligned}$$

As with the specification of agreement with topical subjects or objects in Kinyarwanda/Kirundi given in Chapter 5, example (10) (page 99), this specifica-

<sup>4</sup>A similar situation obtains in Maithili (Indo-Aryan) (Stump and Yadav 1988, Comrie 2003, Bickel et al. 1999), in which the controller of agreement must be prominent in some sense, but can be a subject, object, possessor, or (in some dialects) the object of a preposition. Dalrymple and Nikolaeva (2005) suggest, following Comrie (2003), that the relevant factor is topicality, though further research has indicated that a more general notion of prominence or contrast may better characterise agreement patterns in Maithili.

<sup>5</sup>The expression [GF–SUBJ] is a **regular expression**, and the minus operator (–) is the complementation operator: the expression [GF–SUBJ] refers to all strings consisting of any grammatical function, but with the string SUBJ removed.

tion uses a **local name**, beginning with a percent sign, which is used when constraints are placed on an f-structure whose grammatical function is uncertain or undetermined. In the first line of these specifications, the local name %t is defined as any f-structure which bears some grammatical function within the clause (GF), or a possessor (optional POSS within the GF), but not a subject (the option SUBJ is removed); this f-structure will be the controller of agreement. The constraints in the second, third and fourth line require the f-structure named %t to be first person dative. The final line ensures that %t bears the information structure role of topic. Similar specifications are relevant for affixes marking topicality of arguments with other combinations of person, number, and case.

Itelmen has two sets of secondary agreement suffixes: one for objects, and one for obliques. The oblique agreement suffix is always a marker of topicality: intransitive verbs do not agree with nontopical obliques, and transitive verbs agree with their (topical or nontopical) objects rather than nontopical obliques. For the third person singular oblique agreement suffix, the following specifications are relevant:

(36) Third person singular oblique agreement (Itelmen):

(↑ OBL PERS) = 3  
 (↑ OBL NUM) = SG  
 ((↑ OBL)<sub>σ</sub> DF) = TOPIC

Other combinations of person and number are treated similarly.

Itelmen object agreement is different: since suffixal agreement is always required on the Itelmen verb, object agreement is only optionally an indicator of topicality of the object, though it does indicate that there is no other topical nonsubject (oblique) element in the clause. In other words, the verb shows object agreement if the object is the only possible agreement controller, or if the other potential controllers are nontopical. Constraints associated with the first person singular object agreement suffix are:

(37) First person singular object agreement (Itelmen):

(↑ OBJ PERS) = 1  
 (↑ OBJ NUM) = SG  
 ¬[((↑ OBL)<sub>σ</sub> DF) = TOPIC]  
 (((↑ OBJ)<sub>σ</sub> DF) = TOPIC)

The first two lines of this specification ensure that the object is first person singular. The third line states that there may be no oblique phrase which bears



the topic role in the clause.<sup>6</sup> The fourth line introduces a default specification of topicality for the object in the presence of object agreement.

### 6.3 Conclusion

We have shown that topichood of a nonsubject element can be explicitly indicated by casemarking (as in Persian or Tariana) or agreement (as in Itelmen and possibly Tabassaran) for objects, obliques, possessors, and other nonsubject grammatical functions. The languages we have examined do not require a unique alignment between information structure role and grammatical function, but provide primary evidence for the relevance of topicality in grammatical marking of nonsubjects.

---

<sup>6</sup>The information structure role of other arguments in the clause, including obliques, may be specified by casemarking, agreement, prosody, phrase structure position, or discourse context, as discussed in Chapter 4, Section 4.3.3.



## *Topicality and DOM*

We have seen that some languages treat nonsubject topics specially in terms of grammatical marking: in such languages, topic marking can apply to a variety of nonsubject elements, and any one of a number of grammatical roles can be the target of topic marking. In this chapter we discuss languages in which only topical **objects** are marked, giving rise to DOM. In some of these languages DOM depends on topicality alone, while in others topicality-based DOM works together with semantic factors.

### **7.1 Objects as grammaticalised secondary topics**

Chapter 6 discussed languages in which a whole range of salient nonsubject elements in the clause can bear topical marking: direct objects, some oblique objects and adjuncts, and sometimes even possessors. Crucially, objects are always candidates for marking of this type: we do not know of any language in which nonsubject topic marking is unavailable for objects, and in fact in many languages grammatical marking of topical nonsubject arguments is restricted to objects. This is DOM.

Languages with DOM overtly mark a close association between objects and topics, just as some languages require and mark a close association between subjects and topics. In general, objects — or, more generally, arguments high on the grammatical function hierarchy — tend to be more topical than other nonsubject elements. Croft (1991) observes that “high topicality” is typical of subjects, while objects are characterised by “medium topicality”, in contrast to other grammatical functions. Rude (1986) and Blake (2001:135) also claim that objects tend to be more topical than obliques and adjuncts, which are located low on the hierarchy.

Indeed, as follows from the definition of secondary topic given in Chapter 3, an utterance with a primary and secondary topic conveys a relation that holds between two salient participants. Since the primary topic is closely associated with the subject function (see Chapter 5, Section 5.3), and languages tend to lexicalise important relations between two participants as transitive verbs, we would expect that the secondary topic would often be realised as the second argument of a transitive verb: the object. This means that the direct object is the result of grammaticalisation of the secondary topic in the same way as the subject is usually thought to be the result of the grammaticalisation of the primary topic. In fact, Givón explicitly addresses the grammatical association of the secondary topic function with objecthood and suggests that in historical terms, objects are grammaticalised secondary topics (Givón 1983b, 1984a,b, 1990, 2001).

Sasse (1984) discusses the inherent connection between information structure roles (or, in his terminology, pragmatic functions) and syntax. He argues that subject and object are pragmatically more prominent than other grammatical functions: the canonical pragmatic function of the direct object as a secondary grammatical function associated with the patient role is to identify the pragmatic peak corresponding to “lower-order” (secondary) topicality. The correlation between objecthood and topicality is so strong that in a number of languages, nontopical patients cannot be expressed as syntactic objects, and must undergo incorporation. Sasse shows that this situation is attested in a number of Eastern Cushitic languages (see Næss 2007 for similar observations).

Conversely, in some languages topical status triggers the promotion of nonobject arguments to the object role, often via applicativisation (see Peterson 2007 and references therein, and the discussion of Upper Necaxa Totonac in Chapter 9, Section 9.4.3). For instance, Rude (1986) shows that in Nez Perce (Sahaptian) a number of nonobject grammatical functions can undergo promotion to object. This is characteristic of allative, associative and ablative obliques. In example (1a), ‘river’ is an oblique argument with allative case, while example (1b) illustrates an alternative encoding of this argument as a direct object marked with the object marker *-ne*; here the verb hosts the applicative affix (“allative voice”), signalling transitivity.

- (1) a. kaa hi-tèem'ik=se pik'úun-x  
 and 3Nom-go.down-Asp river-All  
 'And she went down to the river.'
- b. kaa kúus-ne pée-x'yuu-ye  
 and water-Obj 3.Tr-go-All-Asp  
 'And he went to the water.' (Rude 1986:139)

Rude argues, based on a textual study, that “promoted”/applied objects are more topical than non-promoted obliques. In fact, he claims that “the Nez Perce direct object is a kind of secondary topic” (Rude 1986:148), although his definition of secondary topic is not actually provided. Rude’s operational criteria for topicality are different from the presuppositional approach we employ here: for him, topicality is a gradient discourse-related notion and can be “measured” in terms of referential distance and persistence in discourse. Nevertheless, the two approaches lead to roughly similar results when it comes to the analysis of narrative texts. A recurrent referent repeatedly mentioned in the previous discourse is likely to be salient for the speaker and the addressee and therefore to be topical in our sense as well. In the remainder of this chapter, we will discuss languages where topicality marking is restricted to objects, parallel to the languages where it is restricted to subjects, as discussed in Chapter 5.

## 7.2 Agreement with topical objects: Tundra Nenets

In Tundra Nenets<sup>1</sup> (Uralic), subject agreement is obligatory and references both person and number features. Object agreement is optional, and references the number but not the person of the accusative object. Thus, intransitive verbs agree with the subject, while transitive verbs either agree with the subject alone, or with both the subject and the object. The object marker for the singular object is always phonologically null, and the marker for singular objects is a portmanteau morpheme referring both to the subject and the object. In further glosses, object agreement verbs are glossed simply as Obj, without indicating the object marker specifically.

Agreeing and nonagreeing objects have different information structure roles. If example (2d) is understood as an answer to (2a), the whole clause constitutes the focus domain. If it is understood as an answer to (2b), the focus domain includes the verb and the object, but excludes the (topical) subject. If it is

<sup>1</sup>The Tundra Nenets data were collected by the second author during fieldwork supported by the Endangered Languages Documentation Programme, SOAS, London, as well as a grant from the Academy of Finland (project number 125225).

understood as an answer to (2c), the object corresponds to narrow focus. In all of these contexts, object agreement is disallowed.

- (2) a. What happened?  
 b. What did a/the man do?  
 c. What did a/the man kill?  
 d. xasawa ti-m            xada°            / \*xada°da  
    man    reindeer-Acc kill.3SgSubj    kill.Obj.3SgSubj  
    ‘A/the man killed a/the reindeer.’

Similarly, in (3a) the question word *xíbya-m* ‘whom’ is focused, and object agreement is impossible. Example (3b) is understood as an answer to (3a), with the object *Peter* in narrow focus: again, object agreement is ungrammatical.

- (3) a. Wanya xíbya-m ladə°            / \*ladə°da  
    John    who-Acc hit.3SgSubj    hit.Obj.3SgSubj  
    ‘Whom did John hit?’  
 b. Wanya Pyetya-m ladə°            / \*ladə°da  
    John    Peter-Acc hit.3SgSubj    hit.Obj.3SgSubj  
    ‘John hit Peter.’

In contrast, agreement must be present when the previous context establishes a topical role for the object. Interpreted as answers to the question in (4a), the object of *kill* in (4b) and (4c) is the secondary topic, and the verb must show object agreement. Thus, Nenets is similar to the languages discussed in Chapter 6 in that agreement depends on topichood; it is syntactically more constrained, however, in that only topical objects control secondary agreement, and not other nonsubject arguments.

- (4) a. What did a/the man do to the/a reindeer?  
 b. xasawa ti-m            xada°da            / \*xada°  
    man    reindeer-Acc kill.Obj.3SgSubj    kill.3SgSubj  
    ‘A/the man killed a/the reindeer.’  
 c. xada°da  
    kill.Obj.3SgSubj  
    ‘He killed it.’

As shown in (4c), the topical subject and object need not be overtly expressed. As noted in Chapter 3, topical arguments are often discourse-old, and tend to receive reduced expression as a pronoun or null element; in fact, a topical third person pronoun is normally omitted unless it is contrastively stressed. When there is no object pronoun, the agreement affix is the only overt expression of the object.

There is no difference in the behaviour of objects with different semantic features such as definiteness or animacy with respect to object agreement. Nonreferential objects do not trigger agreement, but this follows from the general condition that topics must be referential, as noted in Chapter 3. Third person pronouns behave like lexical nouns, as shown in (5), which is grammatical without object agreement if the object is in focus, for example as an answer to the question *Who did John hit?* or *What did John do?*. If it is construed as an answer to the question *What did John do to him?*, so that the object has the secondary topic role, object agreement is obligatory.

- (5) Wanya syita ladə° / ladə°da  
 John he.Acc hit.3SgSubj hit.Obj.3SgSubj  
 ‘John hit him.’

Note that in some respects, object agreement in Nenets resembles that of Chicheŵa, as analysed by Bresnan and Mchombo (1987) (see Chapter 2, Section 2.6). As in Nenets, Chicheŵa object agreement is optional, and correlates with the topicality of the object. However, a closer look reveals that Nenets object agreement is *grammatical* agreement, whereas Chicheŵa involves what Bresnan and Mchombo (1987) call *anaphoric agreement*, or pronominal incorporation. Bresnan and Mchombo (1987) point out that in languages with incorporated pronominal objects, the verb cannot govern the case of the full noun phrases that are anaphorically linked to the incorporated pronominals, since these full noun phrases are not arguments of the verb. This, then, allows us to distinguish between anaphoric and grammatical agreement. Crucially, objects in Nenets must appear in accusative case, even if object agreement is present on the verb, as shown in (5). This shows that object noun phrases in Nenets are governed by the verb, and that the object agreement affixes represent grammatical agreement and not pronominal incorporation.

In (6), narrow focus is associated with the locative phrase, the topic referent is John, and the secondary topic referent is Peter; again, object agreement is obligatory:

- (6) a. Where did John hit Peter?  
 b. Wanya Pyetya-m pedara-x<sup>o</sup>na ladə<sup>o</sup>da / \*ladə<sup>o</sup>  
 John Peter-Acc forest-Loc hit.Obj.3SgSubj hit.3SgSubj  
 ‘John hit Peter in the forest.’

Nenets allows nontopical (focus) subjects in transitive clauses; they are not morphosyntactically marked, but must bear nuclear stress. If the object is topical, the verb shows object agreement:

- (7) a. xíbya ti-m xada<sup>o</sup> / xada<sup>o</sup>da  
 who reindeer-Acc kill.3SgSubj kill.Obj.3SgSubj  
 ‘Who killed a/the reindeer?’  
 b. Wanya ti-m xada<sup>o</sup> / xada<sup>o</sup>da  
 John reindeer-Acc kill.3SgSubj kill.Obj.3SgSubj  
 ‘JOHN killed a/the reindeer.’

There is no subject topic in these examples. This indicates that object agreement is not associated only with the secondary topic function. In (7b), the agreeing object is the primary topic and is likely to have a definite interpretation. In other words, although in many cases the agreeing object bears the secondary topic role, the function of object agreement is broader: it can simply indicate the topicality of the object.

However, there are certain semantic restrictions on agreeing objects in Tundra Nenets. Objects with indefinite determiners never trigger agreement:

- (8) xurkaxəw<sup>o</sup> pyíryibtya-m pyúr<sup>o</sup>ŋa / \*pyúr<sup>o</sup>ŋada  
 some girl-Acc look.3SgSubj look.Obj.3SgSubj  
 ‘He is looking for some girl.’

Additionally, first and second person singular and plural object pronouns in Nenets behave differently from all other objects in that they do not trigger agreement, no matter what their information structure role. Example (9) is ungrammatical with object agreement, even construed as an answer to the question *What did John do to you/me?*, a context in which the object is associated with the secondary topic function.

- (9) Wanya syiqm<sup>o</sup>/syit<sup>o</sup> ladə<sup>o</sup> / \*ladə<sup>o</sup>da  
 John I.Acc/you.Acc hit.3SgSubj hit.Obj.3SgSubj  
 ‘John hit me/you.’



For first and second person pronominal objects, then, patterns of agreement do not depend on information structure. Instead, they are defined in terms of the referential status of these objects: first and second person objects never trigger agreement. We return to this point in Chapter 10.<sup>2</sup>

We now turn to the question of the grammatical function of agreeing vs. non-agreeing objects, as determined by their syntactic behaviour: we find no behavioural differences that distinguish them, and we analyse both as the primary object, LFG's OBJ. Objects in Nenets have a number of syntactic properties that distinguish them from other grammatical functions. They can be promoted to subject in the passive. Subjects and objects are the only two grammatical functions that can be relativised using the participial strategy; all other grammatical functions must be relativised by means of another verbal form, the action nominal. These tests, however, will not help us in establishing syntactic differences between marked and unmarked objects, since there is no object agreement in the resulting construction.

Another object property relates to control structures. Nenets has a number of complement-taking verbs which take a dependent null-subject clause headed by a so-called converb (either the modal converb or the purposive converb). The dependent subject must be interpreted as coreferential with the matrix object: this is object control. Both agreeing and nonagreeing objects can control the dependent subject.

- (10) a. nyísyá-da nyú-m-ta xanye°  
 father-3Sg son-Acc-3Sg hunt.Mod.Conv  
 toxola°/toxola°da  
 teach.3SgSubj/teach.Obj.3SgSubj  
 'The father taught his son to hunt.'
- b. xæ-wəncy° nya-m-ta  
 leave-Purp.Conv friend-Acc-3Sg  
 xəlkad°tampyi/xəlkad°tampyida  
 persuade.3SgSubj/persuade.Obj.3SgSubj  
 'He is persuading his friend to leave.'

Additionally, both agreeing and non-agreeing objects can serve as the antecedent of a possessive reflexive, provided the antecedent linearly precedes the reflexive.

- (11) Pyetya Masha-m pida mya-k°na-nta ladə° / ladə°da  
 Peter Mary-Acc she yurt-Loc-3Sg hit.3SgSubj hit.Obj.3SgSubj  
 'Peter hit Mary<sub>i</sub> in her<sub>i</sub> yurt.'

<sup>2</sup>In some varieties of Nenets, third person pronouns behave like first and second person pronouns, and do not trigger agreement.

Moreover, Nenets agreeing and non-agreeing objects do not show any obvious positional difference. Nenets is a fairly strictly subject-initial verb-final language, but word order is otherwise relatively free. Example (12a) demonstrates that both types of objects can appear immediately before the verb, while in (12b) both types of objects are separated from the verb by the oblique element ‘in the forest’.

- (12) a. nyísya-da pedara-x°na wenyako-m ladə° / ladə°da  
 father-3Sg forest-Loc dog-Acc hit.3SgSubj hit.Obj.3SgSubj  
 ‘His father hit a/the dog in the forest.’
- b. nyísya-da wenyako-m pedara-x°na ladə° / ladə°da  
 father-3Sg dog-Acc forest-Loc hit.3SgSubj hit.Obj.3SgSubj  
 ‘His father hit a/the dog in the forest.’

We therefore suggest that marked and unmarked objects realise the same grammatical function: the object. Nenets is a language with only one object function, the OBJ function of LFG (see Chapter 2, Section 2.2). As we might expect, then, Nenets has no double object constructions; the goal argument of verbs such as ‘give’ is invariably expressed by a dative-marked oblique:<sup>3</sup>

- (13) Petya Masha-n°h ti-m myiqqa / myiqqada  
 Peter Masha-Dat reindeer-Acc give.3SgSubj give.Obj.3SgSubj  
 ‘Peter gave Masha a/the reindeer.’

Dative obliques never trigger agreement and do not have other syntactic properties of objects: they do not passivise, do not participate in control constructions and are relativised by means of a different strategy.

In formal terms, third person topical object agreement in Nenets is associated with the following constraint:

- (14) Agreement with third person topical objects:

$$(\uparrow \text{ OBJ PERS}) = 3$$

$$((\uparrow \text{ OBJ})_{\sigma} \text{ DF}) = \text{TOPIC}$$

This is similar to the constraint for topical subject agreement, given in (18) of Chapter 5, except that the constraint requires the object rather than the subject to be topical. As a result of this specification, the semantic structure contributed by the object is associated with the information structure role of topic.

<sup>3</sup>In Chapter 8, we will see that there are languages that have more than one object function, the primary OBJ and the secondary OBJ<sub>θ</sub> of LFG, but nevertheless do not have a double object construction; Ostyak, an Uralic language related to Nenets, is one such language. We discuss these patterns further in Chapter 9.

Verbs which agree only with the subject are not associated with this constraint, since they do not require their object to be topical (recall that first and second person objects do not control agreement, and may or may not be topical).

### 7.3 Casemarking of topical objects

In Tigre (Semitic) and Dolakha Newar (Tibeto-Burman), casemarked and non-casemarked objects have different information structure roles: marked objects are topical, while unmarked objects are nontopical. We know of no behavioural differences between marked and unmarked objects, and we analyse both as primary objects, LFG's OBJ. Unlike Nenets, both languages have a double object construction, which we discuss in Chapter 9; here we restrict attention to the behaviour of monotransitive objects, marked and unmarked, showing that they bear different information structure roles, but correspond to the same grammatical function.

#### 7.3.1 Tigre

The primary description of the Tigre (Semitic) data comes from Jake (1980).

Tigre objects can be marked by the preposition *?igil*, and can trigger object agreement in gender. Agreement is determined by definiteness: definite objects always trigger agreement, while indefinites do not. This distribution is not affected by information structure. Here we are primarily interested in prepositional marking of objects, which is optional on definite objects of monotransitive verbs, but incompatible with indefinite objects (animacy plays no role). For monotransitive verbs, this results in the following possibilities for object marking: (i) casemarked definite objects that trigger agreement; (ii) noncasemarked definite objects that trigger agreement, and (iii) noncasemarked indefinite objects that do not trigger agreement.<sup>4</sup> In Chapter 9, we return to a discussion of casemarking and agreement in double object constructions in Tigre.

The objects in examples (15) and (16) are definite, and must agree with the verb. The preposition is allowed but not required for the definite object in (15), and disallowed with the indefinite object in (16). We are interested in the optionality demonstrated in (15), and the conditions under which definite objects must be preceded by the preposition *?igil*.

<sup>4</sup>Jake (1980) notes that the causee argument displays the same distribution of prepositional marking as patient/theme objects: indefinite causees are unmarked, while definite causees are either marked or unmarked. However she argues that the causee differs from the patient/theme objects in a number of other syntactic properties, from which she concludes that it corresponds to a different grammatical function. Since the data on causatives are complex, we leave this question open.

- (15) ḥasāmā      ?ittā (ʔigil) la ḥis'an nadʔayu      / \*nadʔā  
 Hasama.Masc to.her Prep the boy sent.3Masc.3Masc sent.3Masc  
 'Hasama sent the boy to her.' (Jake 1980:72)

- (16) Lilat      (\*ʔigil) waraqat katbat(\*tā)  
 Lilet.Fem Prep letter.Fem wrote.3Fem(\*3Fem)  
 'Lilet wrote a letter.' (Jake 1980:73)

Jake (1980) is mainly concerned with an explication of conditions on verb agreement, and she does not provide a detailed characterisation of when the preposition *ʔigil* is used on definite objects. However, an analysis of the narrative texts published in Raz (1983) reveals that such objects must be topical. A definite object introduced into the discourse for the first time remains unmarked, even if it is highly identifiable; we suggest that this is because there is no pragmatically presupposed relationship between its referent and another salient referent, and so the object referent is not a secondary topic. This is shown in examples (17) and (18).<sup>5</sup>

- (17) a. wa kəʔanna ʔət ləwaddu mədər kanaʔan warsaw  
 and like.this while doing land Canaan inherited.3PI  
 'And while living like this, they inherited the land of Canaan.'  
 (Raz 1983:111)
- b. gis wagabilye mən ʔəde farʔon ʔafgər  
 go and.my.people from hand Pharaoh bring.out  
 'Go and free my people from the hands of Pharaoh.' (Raz 1983:109)

The objects in these examples are highly definite: in (17a) the object is a proper name ('Canaan') and in (17b) it is a possessed definite phrase ('my people'). However, there is no presupposed pragmatic relationship between the subject and the object referent established prior to the time of the utterance. Example (17a) is the last sentence of a text telling the story of Moses and the people of Israel, but there is no previous mention of the land of Canaan in the text. Example (17b) is direct speech: God is addressing Moses. Again there is no established pragmatic relation between the subject (Moses) and the referent of 'my people', since it is the first time God addresses Moses with this request and it is rather unexpected for Moses.

<sup>5</sup>We have followed Raz's transcription conventions for the examples taken from his work, which differ from the conventions used by Jake. Raz transcribes the object-marking preposition as ʔəgəl.

On the other hand, discourse-old definite objects are likely to be topical, and are often marked by the preposition *?igil*. The beginning of the same story describes how the mother of Moses made a chest, put her baby in the chest, and put the chest in the Nile. The chest was then found by Pharaoh's daughter. At that stage of the story the NP denoting the referent 'chest' is definite and identifiable by the interlocutors, since it has been mentioned in the previous discourse. Example (18) describes the first time the Pharaoh's daughter sees the chest; there is no presupposed pragmatic relationship between her and the chest, and the object remains unmarked.

- (18) wa ?əttu mən rayim lasanduqat salsala ?ətta ma-y kərit  
 and there from afar the.chest reed in.the water placed  
 rə?etta  
 saw.3Fem.it  
 'And there she saw, from afar, the reed chest placed in the water.'  
 (Raz 1983:108)

Pharaoh's daughter then sends her maidservants to retrieve the chest. When sentence (19) is produced, the interlocutors have a mental representation of a certain relation that holds between the maidservants and the chest, since the narrator has already made it clear that the servants were sent for the chest. The sentence provides new information about the relationship that holds between the maidservants (primary topic) and the chest (secondary topic), and can be paraphrased as follows: 'what the maidservants then did to the chest is: they opened it'. The object NP 'the chest' is prepositionally marked.

- (19) wa lawašāyfa ?əgəl lasanduqat kfət-kəmsal-?abalaya gōna  
 and her.maid.servants Prep the.chest when-opened.3Pl.it a.child  
 bakke rakbaya  
 crying found.3Pl  
 'And when her maidservants opened the chest, they found [in it] a child  
 crying.'  
 (Raz 1983:108)

Note that there is no prepositional marking on the indefinite object 'child' in the second clause.

Similarly, example (20), taken from a tale about two friends, a cat and a dog, updates the nature of the relationship that holds between two highly salient participants ('the cat deceived the dog').

- (20) da?am dəmmu ... ?əgəl kaləb wə?ul talmat ?əttu  
 but cat Prep dog deliberately she.deceived him  
 'But the cat ... deliberately deceived the dog.'  
 (Raz 1983:104)

Here the subject ('cat') and the object ('dog') are construed as the primary and the secondary topic, respectively. The secondary topic object exhibits prepositional topic marking.

Thus, casemarked objects in Tigre are always topical. The examples we have seen indicate that the casemarked object is the secondary topic, but we have no examples where the subject is in focus, so we do not know if the casemarker appears on primary topic objects as well. Only a subset of topics is marked, namely, definite topics; indefinite topical objects are not marked. Tigre object marking involves, then, a combination of semantic and information structural factors.

The next question is whether casemarked and noncasemarked objects correspond to the same grammatical function. Jake (1980) shows that casemarked and noncasemarked definite objects of monotransitive verbs behave identically in obligatorily triggering agreement, and she does not discuss any other behavioural differences between the two types of monotransitive objects. In the absence of any evidence of behavioural differences, we believe that case-marked and noncasemarked objects of monotransitive verbs correspond to the same grammatical function: the (primary) object, LFG's OBJ.

The formal treatment of topical object casemarking for monotransitives is straightforward:

(21) Casemarking of topical objects, Tigre:

$$\begin{aligned} &(\text{OBJ } \uparrow) \\ &(\uparrow \text{ DEF}) = + \\ &(\uparrow_{\sigma} \text{ DF}) = \text{TOPIC} \end{aligned}$$

These constraints require the marked argument to be an object, to be definite, and to play the role of topic at information structure. In fact, the treatment of casemarking is more complicated than this when we take into account patterns of casemarking with ditransitive verbs; we return to a discussion of ditransitive marking in Tigre in Chapter 9.

Kifle (2007) discusses the closely related language Tigrinya, and shows that, although patterns of casemarking and agreement in Tigrinya are similar in very broad terms to Tigre, the two languages differ in interesting ways. In particular, it is casemarking in Tigrinya that depends on definiteness, while agreement patterns are determined by information structure role: in this way, Tigrinya is, in a sense, "opposite" to Tigre with respect to the roles of agreement and casemarking in signalling semantic information and information structure role. As in Nenets, agreement depends on topicality for transitive objects. Additionally, Tigrinya has an applicative construction in which the applied argument must be topical and verb agreement with the applied object is required; this follows the general tendency for applicativisation to be triggered by topical status of

the applied argument, as discussed for Nez Perce in Section 7.1 of this chapter. Interestingly, the applied argument cannot become the subject of a passive sentence, although either the recipient or the theme of an underived ditransitive verb can passivise; Kifle (2007) analyses this as a problem for the alignment between information structure roles and grammatical functions which we propose, but we believe that it actually indicates a need for refinement of the general theory of applicatives and voice alternations, and in particular Bresnan and Moshi's (1990) formal theory of object asymmetries; the Tigrinya data are otherwise unproblematic for our view.

### 7.3.2 Dolakha Newar

Object casemarking in Dolakha Newar (Tibeto-Burman), as described by Genetti (1994, 1997, 2007), shows essentially the same properties as Tigre, except that definiteness does not play a role. Objects are either unmarked (hence in absolutive case) or marked with the suffix *-ta*, termed "dative" by Genetti. The suffix will be glossed here as Obj, although it should be noted that the same marking occurs on experiencer subjects. We return in Chapter 9, Section 9.4.1 to patterns of casemarking for with ditransitive verbs.

Objects of monotransitive verbs are either marked or unmarked. According to Genetti (2007:113), the object is casemarked if (i) the referent is human and "given" in the discourse, or (ii) the referent is nonhuman but animate, and "occurs in a clause crucial to the resolution of a narrative plot". This means that marked objects denoting human referents tend to be discourse-old. In example (22), the child was mentioned in the immediately preceding sentence in the discourse ('Then they had one small son at that time.'). The utterance makes an assertion about the relationship that holds between two highly topical referents under discussion, the parent and the child.

- (22)  $\bar{a}m$   $\mu\bar{a}c\bar{a}-ta$   $\bar{b}\bar{a}b\bar{u}-ri-n$   $\mu\bar{a}c\bar{a}$   $j\bar{u}-e-l\bar{a}gin$   $\mu\bar{u}r\bar{y}\bar{a}-ku$   
 that child-Obj father-Ind-Erg child be-NMLZ-because lap-Loc  
 ta-ene  
 put-Part  
 'Because he was a child, the father put the child on his lap.'  
 (Genetti 2007:115)

Notice, however, that discourse givenness is not actually a necessary condition for object marking. Genetti provides evidence that in some cases the referent of a marked object is not mentioned in the previous text, but is "accessible through the invocation of a schema". Example (23) is taken from a text about a crown prince; in Dolakha Newar culture it is generally assumed that crown princes have wives, so the referent of the phrase 'your legal wife' is accessible

and its existence is presupposed. Recall from Chapter 3, Section 3.2.2 that pragmatic presupposition of existence is a necessary property of topics, while discourse givenness is not.

- (23) chana bihaiti-ta-uri chin chē-ku ta-ina  
 2Sg.Gen legal.wife-Obj-Top 2Sg.Erg house-Loc put-Fut  
 ‘You will put your legal wife in the house.’ (Genetti 1997:48)

On the other hand, not all human objects are casemarked. They typically remain unmarked when first introduced into the text, so humanness is not a sufficient condition for object marking.

For nonhumans, too, object marking indicates that the referent is highly salient, since such objects occur at a climax in the narrative. Example (24) appears “when the manipulation of the animals is crucial to plot resolution” (Genetti 1994:51). The speaker has related the son’s plan to release the calf, and the release of the calf described here is the culmination of this plan.

- (24) kae-uri-n tapakka sācā-ta phen-ju  
 son-Indef-Erg all.at.once calf-Obj release-3Sg.Past  
 ‘Then the son suddenly released the calf.’ (Genetti 2007:114)

In our terms, this means that the object in (24) is topical.

Genetti also provides several examples of casemarked inanimate objects, although she notes that they are rare:

- (25) a. ota ultā yeŋ-an keŋ-gu ju-en con-a  
 this.Obj translate do-Part show-NMLZ be-Part stay-3Sg.Past  
 ka  
 Associative  
 ‘It turns out she translates this, then shows it to people.’  
 (Genetti 1994:51)
- b. āu luŋmā tuphi-ta hāti yer-eu?  
 now mortar broom-Obj what do-3Pl.Fut  
 ‘Now what will he do with the mortar and the broom?’  
 (Genetti 1994:114)

In example (25a), the object marker has fused with the pronoun. In both examples the object is under discussion, as can be seen by examination of the previous context. For example, (25b) is taken from a narrative text. The immediately preceding discourse is: ‘One of them quickly brought in a mortar and a broom. And put them there. What was born? They said: A mortar and



broom were born. When they said that the king stayed silent.’ In this context, (25b) describes a pragmatically highly salient relationship between the king and the mortar and broom, where the coordinated object NP is the secondary topic. These examples show that inanimate objects may be casemarked if they are topical.

Thus, casemarking of patient/theme objects cannot be unambiguously described in terms of definiteness, discourse-givenness or animacy. Instead, the presence of the casemarker requires a certain degree of pragmatic salience: the referents of the object and the subject must stand in a certain presupposed relationship established in the context or based on world knowledge. The assertion associated with the sentence where the marked object appears is meant to update the addressee’s knowledge about this relationship. Notice that object marking can co-occur with the topic marker, as shown in example (23) as well as in (26):

- (26) bhut-na janta wā guli khyāṅ-an tar-ai  
ghost-Erg 1SgObj Top how.much scare-Part put-3Sg.Pres  
‘Ghosts scare me so much.’ (Genetti 2007:300)

It is not clear on Genetti’s account whether the topic marker in either of these examples carries an additional meaning (for instance, contrastivity).

Genetti (2007:315) shows that objects in Dolakha Newar can be distinguished from subjects by means of a number of tests: objects do not trigger agreement, do not serve as antecedents of regular reflexives, and do not participate in control constructions or certain types of relativisation. The tests do not distinguish between marked and unmarked objects, however, and she concludes that there is no behavioural syntactic difference between marked and unmarked objects: in our terms, they are both primary objects, LFG’s OBJ. We adopt a formal treatment of object marking in Newar that is similar to the one proposed above for Tigre, except that definiteness does not play a role in marking:

- (27) Casemarking of topical objects, Newar:

(OBJ ↑)  
(↑<sub>σ</sub> DF) = TOPIC

These constraints require the marked argument to be an object and to fill the topic role at information structure.

## 7.4 Conclusion

In the languages we have examined in this chapter, marked and unmarked objects of monotransitive verbs correspond to the same grammatical function:

both are primary objects, the OBJ function of LFG. We base this claim on the lack of observed behavioural differences between them. Of course, further research on these languages may reveal hitherto undiscovered behavioural differences between marked and unmarked objects, which would necessitate their reclassification along the lines to be discussed in Chapter 8. There, we will see that other languages exhibit a different pattern: the marked object of mono-transitive verbs can be shown to have different grammatical properties from the unmarked object.

## *Primary and secondary objecthood and DOM*

In some languages, differential object marking correlates with a difference in grammatical function: marked and unmarked objects have different syntactic behaviour, with marked objects exhibiting more properties of core grammatical functions than unmarked objects. Here we discuss languages where marked (topical) objects are grammatically distinguished from unmarked (nontopical) objects. We concentrate on marking patterns with monotransitive verbs in this chapter; Chapter 9 discusses marking and alignment patterns for trivalent verbs, with particular attention to languages with ditransitive constructions.

### **8.1 Grammatical marking and grammatical function**

There are two basic patterns of interaction between the grammatical marking of nonsubject topics (DOM) and grammatical objecthood. In languages of the first type, a difference in object marking does not correlate with a difference in grammatical function. Such patterns are not surprising in the context of traditional theories of argument mapping, discussed in Chapter 2, Section 2.1.2, which define possible relations between grammatical functions and semantic roles as they are represented at argument structure. On this view, we would expect marked and unmarked objects which correspond to the same semantic role to be mapped to the same grammatical function, and there is no expectation that information structure role could affect argument mapping.

The languages discussed in Chapter 7 are of this type. In these languages, grammatically marked and unmarked objects do not display behavioural syntactic differences; grammatical marking correlates with information structure differences — topical vs. nontopical arguments — and does not reflect a dif-

ference in grammatical function. In these languages, DOM is defined by information structural considerations, not grammatical function.

In languages of the second type, grammatical marking of the object signals a **difference in grammatical function** and a concomitant difference in grammatical behaviour. This complicates the theory of argument mapping, but makes the relation of f-structure to information structure more clear, since different information structure roles correspond to different grammatical functions. We demonstrate in the following that in languages of the second type, when a predicate takes a single nonsubject nominal argument, it can be encoded as either a topical, marked OBJ, or a nontopical, unmarked OBJ<sub>θ</sub>. This gives rise to patterns of DOM which reflect a difference in syntactic status between marked and unmarked objects.

Languages of this type are Ostyak, Mongolian, Chatino, and Hindi, to be discussed in this chapter. In these languages, marked and unmarked objects show different behavioural profiles. Our proposal allows, and indeed requires, the restatement of generalisations about object agreement and casemarking in these languages in purely syntactic terms: the agreeing or casemarked object is the primary object OBJ, and the nonagreeing or noncasemarked object is the secondary object, a member of the family of grammatical functions OBJ<sub>θ</sub>. OBJ and OBJ<sub>θ</sub> require different patterns of casemarking or agreement, and hence we have DOM. Languages of this type are in essence just like many other languages in defining agreement and casemarking patterns in purely syntactic terms, by reference to grammatical functions rather than information structure. Their distinguishing property is the tight linkage between information structure and f-structure: marked/primary objects (OBJ) are (secondary) topics, while unmarked secondary or restricted objects (OBJ<sub>θ</sub>) are nontopical. Though it appears that secondary agreement in Ostyak and object casemarking in Mongolian, Chatino, and Hindi are determined by information structure, as in the languages examined so far, this is actually a side-effect of the strong relation between information structure role and grammatical function.

## 8.2 Object agreement and grammatical function: Ostyak

Northern Ostyak, also called Northern Khanty, is a Uralic language. The data here reflect the dialect of Obdorsk and come from the fieldwork of the second author.

As in the related language Nenets, subject agreement in Ostyak is obligatory, while object agreement is optional (Nikolaeva 1999, 2001). Intransitive verbs agree with the subject in person and number (1a). Transitive verbs agree either with the subject alone (1b), or with the subject and the object, as in (1c)-(1e). Subject and object pronouns may but need not appear. As in Nenets, object

agreement forms indicate the number (but not the person) of the object. There is no verb agreement other than with subject and object.

- (1) a. (ma) je:lən o:məs-l-əm  
 I at.home sit-Pres-1SgSubj  
 'I am sitting at home.'
- b. (ma) tam kalaŋ we:l-s-əm  
 I this reindeer kill-Past-1SgSubj  
 'I killed this reindeer.'
- c. (ma) tam kalaŋ we:l-s-∅-e:m  
 I this reindeer kill-Past-SgObj-1SgSubj  
 'I killed this reindeer.'
- d. (ma) tam kalaŋ-ət we:l-sə-l-am  
 I these reindeer-Pl kill-Past-PlObj-1SgSubj  
 'I killed these reindeer.'
- e. (ma) tam kalaŋ-ŋən we:l-sə-ŋil-am  
 I these reindeer-Du kill-Past-Du.Obj-1SgSubj  
 'I killed these (two) reindeer.'

Subject markers differ in these forms. For example, the first person singular subject marker in the absence of object agreement is *-əm*, as in (1a) and (1b); with the singular object it is *-e:m* (1c), and with the dual or plural object it is *-am*, as in (1d) and (1e). Similar distinctions obtain for the whole paradigm. As in Nenets, the subject marker with singular objects, such as *-e:m*, is a portman-teau morpheme referring both to the subject and the object, and we gloss object agreement verbs as Obj, without indicating the object marker specifically.

As shown by Nikolaeva (1999, 2001), traditional descriptions of *Ostyak* (Rédei 1965, Honti 1984, and others), which suggest that agreement is conditioned by definiteness of the object, are incorrect: both definite and indefinite objects may but need not trigger agreement. Instead, agreement correlates with information structure role. Objects that trigger agreement are topical, while nonagreeing objects are nontopical and share a cluster of semantic and pragmatic properties that are associated with narrow or wide focus.

We first examine nonagreeing objects. In object questions and answers (narrow focus structures), agreement must be absent, even if the object is definite:

- (2) a. u:r-na    mati kalaŋ    we:l-əs                    /  
 forest-Loc which reindeer kill-Past.3SgSubj  
 \*we:l-s-əlli  
 kill-Past-Obj.3SgSubj  
 ‘Which reindeer did he kill in the forest?’
- b. u:r-na    tam kalaŋ    we:l-əs                    / \*we:l-s-əlli  
 forest-Loc this reindeer kill-Past.3SgSubj    kill-Past-Obj.3SgSubj  
 ‘He killed this reindeer in the forest.’

Example (2b) is also an appropriate answer to the question *What did he do?* (with different prosody). In this case, the object is not the narrow focus, but is a part of the wide focus domain. Both types of objects are nontopical, and the verb agrees with neither.

Objects under the scope of focus items such as *only* or *even* and contrastively focused objects, including first and second person pronouns, also fail to trigger agreement:

- (3) a. tamxatl tup    wul a:n    wa:n-s-əm                    / \*wa:n-s-e:m  
 today    only big cup see-Past-1SgSubj    see-Past-Obj.1SgSubj  
 ‘Today I only saw the/a big cup.’
- b. tamxatl ma-ne:m    jir-əs                    /  
 today    I-Acc    tie.down-Past.3SgSubj  
 \*jir-s-əlli                                    anta naŋ-e:n  
 tie.down-Past-Obj.3SgSubj not you-Acc  
 ‘Today he tied me down, not you.’

It is well known that in many SOV languages there is an immediately pre-verbal position which is reserved for a focused element, independent of its semantic role and grammatical function (Kim 1988, É. Kiss 1995, Butt and King 1996). Ostyak is such a language, with a grammaticalised focus position immediately before the verb. The questioned NP ‘which reindeer’ in (4) must appear in this position.

- (4) a. \*mati kalaŋ    u:r-na    we:l-əs  
 which reindeer forest-Loc kill-Past.3SgSubj  
 ‘Which reindeer did he kill in the forest?’

The same applies to nonagreeing objects that do not correspond to a wh-question word: they are normally impossible to separate from the verb.

- (5) \*tam kalaŋ u:r-na we:l-əs  
 this reindeer forest-Loc kill-Past.3SgSubj  
 'He killed this reindeer in the forest.'

This additionally argues for the focus status of at least some nonagreeing objects, since they are required to appear in focus position.

However, not all nonagreeing objects are in focus, and not all appear in verb-adjacent position. Nonreferential objects never trigger agreement, even if they are not in the focus position and arguably do not bear the focus role. In (5), the focus function is associated with an oblique element (*xalśa* 'where'), which must be immediately preverbal:

- (6) li-ti pil xalśa kas-l-əm / \*kas-l-e:m  
 eat-Part companion where find-Pres-1SgSubj find-Pres-Obj.1SgSubj  
 'Where shall I find a companion to eat with?'

Examples such as (6) illustrate the only circumstance in which non-agreeing objects are separated from the verb by another constituent. Such examples are important because they show that agreement does not directly depend on the position of the object. Both agreeing and nonagreeing objects can be located in a position which is not adjacent to the verb, but only nonreferential nonagreeing objects can appear there.

Given the assumption that topics must be referential, the proper generalisation about Ostyak nonagreeing objects is, then, that they may not be topical. When they appear in the preverbal focus position, they are in focus. When they do not appear in the focus position, they are nonreferential and therefore nontopical. In this case they can be separated from the verb by a focus element.

On the other hand, the position of objects that trigger agreement is fairly free. The object can be separated from the verb by other elements or, in some cases, even appear after it. For referential objects that are not immediately preverbal, object agreement is required:

- (7) Pe:tra mo:jpər u:r-na wa:n-s-əlli / \*wa:nt-əs  
 Peter bear forest-Loc see-Past-Obj.3SgSubj see-Past.3SgSubj  
 'Peter saw the/a bear in the forest.'

Some referential objects appear immediately before the verb, but this only occurs if there is no nonverbal focus constituent, so that either the focus falls on the verb itself or the verb is included in the focus domain. For example, the following sentence can be interpreted in two ways: either as an answer to the question 'What did you do to this reindeer?' (broad predicate focus) or to the question 'You didn't kill this reindeer, did you?' (polarity/verum focus).

- (8) tam kalaŋ we:l-s-e:m  
 this reindeer kill-Past-Obj.1SgSubj  
 ‘I DID kill this reindeer/I KILLED this reindeer.’

Nikolaeva (2001) shows that for some Ostyak verbs object agreement is either highly unlikely or altogether impossible. These are verbs that typically introduce new (nontopical) participants into the discourse. For example, the verb *taj-* ‘have’ very rarely takes object agreement, as confirmed by textual counts. Its meaning is such that in most cases its object cannot be interpreted as topical, cf. the following infelicitous piece of discourse:

- (9) a. What do you do to this knife?/What is the relationship between you and this knife?  
 b. ??‘I have it/this knife.’

Under certain discourse conditions this verb can take object agreement, but such cases are rare. One possible context is the so-called “*verum*” focus.

- (10) a. But you don’t have the/a knife!  
 b. taj-l-e:m!  
 have-Pres-Obj.1SgSubj  
 ‘I do (have it/one)!’

In such examples the focus extends over the verb alone, while the object is topical and therefore triggers agreement.

As discussed in Chapter 5, Section 5.2, the primary topic role is often associated with subject. In Ostyak this association is obligatory, at least in transitive clauses. It follows from this that although agreeing objects are topical, they are not primary topics, since the primary topic role is always associated with the subject. Rather, they must be analysed as secondary topics, as confirmed by the following examples. Clauses with object agreement, as in examples (11b) and (11c), are felicitous as answers to the question in (11a).

- (11) a. What did you do to this reindeer?  
 b. tam kalaŋ we:l-s-e:m / \*we:l-s-əm  
 this reindeer kill-Past-Obj.1SgSubj kill-Past-1SgSubj  
 ‘I killed this reindeer.’  
 c. we:l-s-e:m / \*we:l-s-əm  
 kill-Past-Obj.1SgSubj kill-Past-1SgSubj  
 ‘I killed it.’



This context establishes the secondary topic role of the object in the answer, and nonagreeing objects are not permitted here. In these examples, the verb is in focus. Object agreement is also present when the focus falls on a non-object constituent, to the exclusion of the verb and the object itself.

- (12) *kalaŋ xalśa we:l-s-əlli / \*we:l-əs*  
 reindeer where kill-Past-Obj.3SgSubj kill-Past.3SgSubj  
 ‘Where did he kill the/a reindeer?’

Sentence (12) exemplifies an argument-focus structure in which the focus is on the oblique constituent ‘where’, the object is topical, and object agreement is obligatory.

As shown in Chapter 3, topical arguments tend to receive reduced expression as a pronoun or null element. Consistent with this claim is the fact that objects that trigger agreement may correspond to a referential null, as in (11c). In fact, a corpus count shows that almost half of the clauses with object agreement have no overt object (Nikolaeva 2001). The use of overt NPs as objects that trigger agreement is a marked option and is usually motivated by the need to disambiguate between several referents. Moreover, as shown by the text analysis reported by Nikolaeva (2001), 83% of the objects that trigger agreement have been mentioned in the previous discourse. Only 17% have not been mentioned previously, although these may refer to entities that are relevant in the speech situation. Further arguments for the topical status of the agreeing object are presented by Nikolaeva (2001).

Nikolaeva (1999) and Dalrymple and Nikolaeva (2005) show that *Ostyak* exhibits clear behavioural syntactic differences distinguishing the agreeing from the nonagreeing object. Nikolaeva (1999:346) provides the table in (13), which indicates that the former exhibits more properties of core grammatical functions than the latter.

- (13) Syntactic properties of subjects and objects in *Ostyak*:

|                                                     | Subject | Agreeing<br>(topical)<br>object | Nonagreeing<br>(nontopical)<br>object |
|-----------------------------------------------------|---------|---------------------------------|---------------------------------------|
| Verbal agreement                                    | +       | +                               | –                                     |
| Control of coreference in action nominal<br>clauses | +       | +                               | –                                     |
| Possessor topicalisation                            | +       | +                               | –                                     |
| Control of possessive reflexivisation               | +       | +                               | –                                     |
| Quantifier float                                    | +       | +                               | –                                     |

For example, the second line of the table indicates that both the subject and the agreeing object can control coreference with the subject of an action nominal

(AN) adverbial clause. Such clauses are introduced by various postpositions, which specify the meaning of the adverbial. In example (14), the main and the embedded clause subjects are coreferential:

- (14) [a:š-e:m<sub>i</sub> tutjux se:wər-m-al sis] luw<sub>i</sub>/∅<sub>i</sub> mosa  
 father-1Sg wood cut-AN-3Sg when he something  
 no:məlm-əs  
 remember-Past.3SgSubj  
 ‘When my father was cutting wood, he remembered something.’

In (15), the embedded clause subject is coreferential with the main clause agreeing object.

- (15) [xul<sub>i</sub> u:n u:l-m-al pata] xul<sub>i</sub>/∅<sub>i</sub> nox an ta:l-s-e:m  
 fish large be-AN-3Sg because fish out not carry-Past-Obj.1SgSubj  
 ‘I didn’t take out the fish because the fish/it was large.’

On the other hand, nonagreeing objects in the main clause cannot control the embedded subject. Example (16) has no object agreement in the main clause and is ungrammatical.

- (16) \*[xul<sub>i</sub> u:n u:l-m-al pata] xul<sub>i</sub>/∅<sub>i</sub> nox an ta:l-s-əm  
 fish large be-AN-3Sg because fish out not carry-Past-1SgSubj  
 ‘I didn’t take out the fish because the fish/it was large.’

Other syntactic tests in Ostyak are discussed in detail by Nikolaeva (1999) and are not repeated here. All of these tests lead to the conclusion that the agreeing object in Ostyak shows more properties of core arguments and has more properties in common with the subject, while the non-agreeing object does not exhibit these properties.

Most previous work on DOM has not provided an account of the correlation between DOM and behavioural syntactic properties distinguishing primary from nonprimary objects. Some transformational analyses establish a correlation between casemarking, agreement, and position of the two types of objects, but do not straightforwardly extend to an account of the other syntactic differences summarised in Table (13) for Ostyak. Instead, we suggest that these generalisations should be expressed by assigning agreeing and non-agreeing objects to different grammatical functions. A similar proposal was made by Butt and King (1996), who propose that unmarked, focused objects that are adjacent to the verb correspond to a different grammatical function from objects that may be marked, and whose position is unconstrained: they analyse the unmarked, verb-adjacent object as the **primary object**, the **OBJ**

in LFG, and the positionally unconstrained object as the **restricted or secondary object** OBJ<sub>θ</sub>. We suggest the opposite distribution: that the grammatically marked, topical object corresponds to OBJ, and the nontopical, unmarked object corresponds to OBJ<sub>θ</sub>, as shown in (17).

(17) Marked and unmarked patient/theme objects in Ostyak:

|                                          | OBJ   | OBJ <sub>θ</sub> |
|------------------------------------------|-------|------------------|
| Marking                                  | yes   | no               |
| Information structure role               | topic | nontopic         |
| Properties of core grammatical functions | yes   | no               |

We believe that our proposal provides a cross-linguistically more satisfactory account, since it accounts for the fact that it is the grammatically marked, topical object rather than the unmarked, nontopical object that displays more properties characteristic of core grammatical functions, as shown for Ostyak in the table in (13).

Our proposal stands in agreement with Butt and King's (1996) view that the distinction between primary and secondary objects is relevant not only for ditransitive constructions, but also for monotransitives: monotransitive verbs in some languages can take a subject and either a primary or a secondary object, and a secondary object may appear even when there is no primary object in the clause. Indeed, although Ostyak distinguishes the primary/topical OBJ from the secondary/nontopical OBJ<sub>θ</sub>, it does not have a double object construction: no verb takes an OBJ and an OBJ<sub>θ</sub> at the same time. With verbs like 'give', either the goal or the theme must appear as an oblique, resulting in two possibilities: if the theme is an object, the goal is a dative oblique, as in (18a); and if the goal is an object, the theme is a locative oblique, as in (18b). We return to an in-depth discussion of triadic verbs and cross-linguistically available ditransitive alignment possibilities in Chapter 9.

- (18) a. ma a:n Pe:tra e:lti ma-s-e:m / ma-s-əm  
 I cup Peter to give-Past-Obj.1SgSubj give-Past-1SgSubj  
 'I gave a/the cup to Peter.'
- b. ma Pe:tra a:n-na ma-s-e:m / \*ma-s-əm  
 I Peter cup-Loc give-Past-Obj.1SgSubj give-Past-1SgSubj  
 'I gave Peter a/the cup.'

Our account meshes well with standard LFG assumptions about permissible relations between semantic roles and grammatical functions, and makes the correct predictions for patterns of argument marking with objects bearing a range of semantic roles. As discussed in Chapter 2, Section 2.2, the primary OBJ of LFG is a **semantically unrestricted** object, which can bear any

of a number of semantic roles. In contrast, Bresnan and Kanerva (1989) define  $OBJ_{\theta}$  as a family of **semantically restricted** objects  $OBJ_{THEME}$ ,  $OBJ_{GOAL}$ , and so on, each of which can be filled only by arguments with particular semantic roles. Crosslinguistically, it is very common for languages to allow only patient/theme arguments as  $OBJ_{\theta}$ : we will refer to this grammatical function by the general term  $OBJ_{PATIENT/THEME}$ . In a language whose inventory of secondary object functions is restricted in this way, the correlation between information structure role and grammatical function may be imperfect, holding only for patient/theme objects and not for objects bearing other semantic roles.

This is the situation in Ostyak. A patient or theme argument can correspond either to a primary  $OBJ$  or a secondary  $OBJ_{PATIENT/THEME}$ , with the choice depending on information structure role. Object arguments with other semantic roles cannot be secondary objects, since Ostyak has only one restricted object function, and does not allow secondary objects associated with other semantic roles such as goal or causee. Instead, goal and causee objects must surface as primary objects and must trigger agreement, whether or not they are topical; the constraint limiting  $OBJ_{\theta}$  to arguments bearing the patient/theme role may not be violated, independent of the information-structure role of the argument. This is further evidence that agreement in Ostyak is defined purely in terms of grammatical functions, and not information structure roles: although mapping rules in Ostyak must be defined by reference to information structure roles as well as semantic roles and grammatical functions, agreement can be straightforwardly characterised in terms of the grammatical function  $OBJ$ .

This is shown by agreement patterns with verbs such as ‘give’, which have a patient and a goal argument: they allow the goal to be an agreeing  $OBJ$ , but not a nonagreeing  $OBJ_{\theta}$ , as shown in (18b). When the goal is an oblique, as in (18a), the patient can be either an agreeing  $OBJ$  or a nonagreeing  $OBJ_{PATIENT}$ .

The same is observed in causatives. Ostyak causatives are productively derived only from intransitive verbs. The causee argument is casemarked in the same way as other objects: it stands in the unmarked nominative case if it is a lexical noun, or in the accusative if it is a personal pronoun. However, the causative construction differs from the regular transitive construction in that it requires object agreement even if the causee object bears the focus function. The sentences in (20) are construed as replies to the question in (19); in the question as well as the answers, the causee is the  $OBJ$  and cannot be the  $OBJ_{\theta}$ , and agreement is required:

- (19) xoj xollə-ptə-s-li?  
 who cry-Caus-Past-Obj.3Sg.Subj  
 ‘Whom did he make cry?’

- (20) a. ma:ne:m xo:llə-ptə-s-li / \*xo:llə-ptə-s  
 I.Acc cry-Caus-Past-Obj.3SgSubj cry-Caus-Past-3SgSubj  
 ‘He made me cry.’
- b. Pe:tra xo:llə-ptə-s-li / \*xo:llə-ptə-s  
 Peter cry-Caus-Past-Obj.3SgSubj cry-Caus-Past-3SgSubj  
 ‘He made Peter cry.’

Thus, object agreement with goal and causee objects does not correlate with information structure: they may not be expressed as OBJ<sub>PATIENT/THEME</sub>, and there is therefore no option to make use of the distinction between OBJ and OBJ<sub>θ</sub> to express their information structure status. In other words, non-patient/non-theme objects must be realised as OBJ, no matter what their information structure role.

Note that non-patient/non-theme objects share all syntactic properties of agreeing patient/theme objects. For example, they can control action nominal clauses:

- (21) [pasa:n e:liti o:məs-t-al sis] Juwan a:n-na ma-s-e:m  
 table at sit-AN-3Sg when John cup-Loc give-Past-Obj.1SgSubj  
 ‘I gave John a cup when he was sitting at the table.’

Other properties listed in Table (13) are also identical, which suggests that non-patient/non-theme objects and agreeing patient/theme objects represent the same grammatical function.

In sum, generalisations regarding verb agreement in *Ostyak* are purely syntactic; the verb agrees with the OBJ, but not with OBJ<sub>PATIENT/THEME</sub>. The agreement affix *-am* specifies that the subject is first person singular, and the object is plural, but does not specify information structure role:

- (22) Agreement specifications for the agreement affix *-am*:
- (↑ SUBJ PERS) = 1  
 (↑ SUBJ NUM) = SG  
 (↑ OBJ NUM) = PL

In *Ostyak*, it is the mapping relations that are complex, referring not only to semantic roles, as is usual in theories of mapping, but also to information structure: topical patients and themes map to OBJ, while nontopical patients and themes are secondary objects OBJ<sub>θ</sub>. Agreement patterns then follow directly from grammatical role.

**8.3 Object casemarking and grammatical function***8.3.1 Mongolian*

Khalkha Mongolian, as described by Guntsetseg (2008), also exhibits differential object casemarking.<sup>1</sup> The accusative *-(i)g* is obligatory on definite NPs (including personal pronouns, proper names and universally quantified objects) and optional on indefinite objects (Guntsetseg 2008). This makes Mongolian DOM different from DOM in Ostyak, which depends purely on information structure role and not on particular semantic features, except for the general requirement for the topic to be referential. DOM in Mongolian is partially dependent on semantic criteria: the accusative marker is required on all definite objects, regardless of their topicality. The possible genesis of such semantic conditions in processes of grammaticalisation is discussed in Chapter 10.

The correlation of marking with specificity in Mongolian is indirect. Whereas the accusatively-marked indefinite object must receive a specific interpretation, the nominative object can be either specific or nonspecific.

- (23) a. Bold neg ohin uns-sen  
       Bold a girl kiss-Past  
       ‘Bold kissed a girl/a certain girl.’
- b. Bold neg ohin-ig uns-sen  
       Bold a girl-Acc kiss-Past  
       ‘Bold kissed a certain girl.’

In general, animate specific indefinites are more likely to be casemarked than indefinites, but this is only a tendency: there is no absolute correlation with animacy. Indefinite inanimates, too, can be either marked or unmarked:

- (24) ter neg zahia(-g) bich-sen  
       he a letter(-Acc) write-Past  
       ‘He wrote a letter.’

Affectedness of the object does seem to be a factor in DOM. Affected objects of verbs whose semantics implies a change of state of the object (e.g. ‘repair’) are more likely to be marked with accusative than objects of intensional verbs such as ‘search’. Again, however, this is only a general tendency: as shown by the examples above, the same verb can take either a nominative or an accusative object.

---

<sup>1</sup>Examples for which the source is not mentioned come from personal communication with Dolgor Guntsetseg.

The nominative is required when the object is clearly in focus. Thus, the object in (24) is zero-marked if the sentence is construed as the answer to the question ‘What did he write?’ or ‘What did he do?’. In such focus contexts, accusative is generally impossible. Example (25) illustrates the same pattern:

- (25) a. What did he do? Or: What did he copy?  
 b. ter neg uguulber(\*-ig) huul-san  
 he a sentence-Acc copy-Past  
 ‘He copied a sentence.’

The accusative is equally impossible on indefinite objects if they have a non-referential interpretation, as shown by the contrast in (26):

- (26) a. bi neg oyutn-ig haij baina. Ter ih uhaantai  
 I a friend-Acc look.for Prog he very clever  
 ‘I am looking for a friend. He is very clever.’  
 b. bi neg oyutan / \*ojutn-ig haij baina. Ter ih uhaantai baih  
 I a friend friend-Acc look.for Prog he very clever be  
 yostoi  
 should  
 ‘I am looking for a friend. He should be very clever.’

In (26b) the object ‘friend’ is nonreferential: the speaker does not imply that the referent of this NP exists. The accusative is strictly ungrammatical in this context.

The question is, then, what triggers the accusative on a referential indefinite object in (26a). We suggest that the relevant factor is topicality. We argued in Chapter 3, Section 3.2.2 that under certain conditions indefinite NPs can be interpreted as topics if they are pragmatically linked to another accessible referent, or if the neighbouring context provides additional information about them. Such context may involve relativisation, as in Persian (Chapter 6, page 114, example 11) and Danish (Chapter 3, page 54, example 7) or a coordinated clause, as in Persian (Chapter 6, page 114, example 12). The same is observed in Mongolian, as seen in example (26a). Here are two more examples:

- (27) a. What did he do?  
 b. ter neg uguulber-ig / \*uguulber huul-san gevch ter ni buruu  
 he a sentence-Acc sentence copy-Past but it Top wrong  
 bai-san  
 be-Past  
 ‘He copied a sentence but it was wrong.’

- (28) a. What did you do yesterday?
- b. bi uchigdur neg nom-ig / \*nom unsh-aad duusga-san. ter ikh  
 I yesterday a book-Acc book read-Conv finish-Past it very  
 sonirkholtoi bai-san  
 interesting be-Past  
 ‘Yesterday I finished a book. It was very interesting.’

In both of these examples, the object is part of the broad predicate focus domain. However, it is topical within this domain, because it is interpreted as pragmatically salient by the speaker who assesses its saliency when planning the subsequent clause, in which the same referent is mentioned. In this context the accusative is obligatory. Without the continuation (‘but it was wrong’ and ‘it was very interesting’) the accusative would be ungrammatical in this context. This means that the accusative marks definite objects and indefinite topical objects. Indefinite nontopical objects remain unmarked.

This conclusion stands in agreement with the suggestion made by Guntsetseg (2008) that the decisive factor determining DOM in Mongolian is discourse prominence, in the sense of topicality. She tested this by means of a questionnaire experiment given to 320 native speakers for translation, consisting of 75 test sentences relating to DOM and 100 filler/control sentences. The questionnaire contained examples where the object is anaphorically cross-referenced in the next clause (e.g. ‘John kissed a girl and she slapped him’) and examples without such a relationship (e.g. ‘John kissed a girl. James didn’t come to school today’). The results show that the indefinite object is more likely to be accusatively marked if the speaker intends to give further information about its referent in the immediate subsequent discourse, where this object is normally cross-referenced by anaphora.

The marked object in Mongolian may be either the primary or the secondary topic. In the presence of a focused subject, topical marking is allowed on the object, as shown in example (29b), where the object is more topical than in (29a):

- (29) a. Hen neg ohin uns-sen be?  
 who a girl kiss-Past Q  
 ‘Who kissed a girl?’ (any girl)
- b. Hen neg ohin-ig uns-sen be?  
 who a girl-Acc kiss-Past Q  
 ‘Who kissed a girl?’ (a certain highly salient girl)



As with the other languages discussed in this chapter, marked and unmarked objects in Mongolian have different syntactic properties, and we assign them to different grammatical functions based on these differences. First, they differ in their ability to combine with the particle *ni*, whose distribution is syntactically governed: it appears only on subjects and marked objects, and not on unmarked objects.

- (30) *bi neg huuhd-ig / \*huuhd ni hav-san*  
 I a child-Acc child NI see-Past  
 ‘I saw a (certain) child (one of the children).’ (Guntsetseg 2008)

On subjects and definite marked objects, *ni* marks a topic which is usually, but not always, contrastive. An example with *ni*-marking on the subject can be found in the second conjunct of (27b). On indefinite marked objects, the particle normally implies a partitive reading (one of a known set).

The second difference is that, although unmarked constituent order in Mongolian is SOV, accusative objects can be fronted to the pre-subject position; this is impossible for nominative objects:

- (31) *neg zahia-g / \*zahia bi bich-sen*  
 a letter-Acc letter I write-Past  
 ‘A letter, I wrote.’ (Guntsetseg 2008)

These differences cannot be explained by assuming that the unmarked/nominative object is required to appear in a particular phrase structure position. Although there is a tendency for the unmarked object to be immediately preverbal (similar to Ostyak, Turkish, Hindi and many other SOV languages, as discussed earlier), Mongolian does not seem to have a dedicated focus position located immediately before the verb. Both the marked and unmarked object can be separated from the verb by other constituents:

- (32) *bi neg ohin(-ig) gudamj-ind zohi-son*  
 I a cat-Acc street-Dat/Loc hit-Past  
 ‘I hit a cat in the street.’ (Guntsetseg 2008)

Unlike in Ostyak, this word order does not necessarily require a nonreferential interpretation of the unmarked object. Further, the focus *wh*-question word need not immediately precede the verb, as shown in example (33):

- (33) *ter haana neg baavgai al-san be?*  
 he where a bear kill-Past Q  
 ‘Where did he kill a bear?’ (Guntsetseg 2008)

Thus, a phrase structure based account cannot account straightforwardly for the different behavioural properties of marked and unmarked objects, as in Mongolian unmarked objects need not appear in a fixed phrase structure position. Instead, we analyse accusative and nominative objects as corresponding to different grammatical functions, the former to OBJ and the latter to OBJ<sub>θ</sub>.

We treat the accusative affix *-ig* as simply marking its argument as an OBJ, in line with Nordlinger's (1998) constructive case theory, described in Chapter 2, Section 2.7. As in Ostyak, rules governing argument mapping are complex, involving semantic role, information structure role, and grammatical function, but marking can be specified simply in terms of grammatical role:

(34) Specification for the accusative affix *-ig*:

(OBJ ↑)

The theory of argument linking which specifies how the arguments of a verb are syntactically realised must take into account the information structure role of an argument in determining whether it appears as an OBJ or an OBJ<sub>θ</sub>, but rules for determining casemarking are very simple, referring only to the grammatical function of the marked argument.

### 8.3.2 *Chatino*

In Chatino (Zapotecan), described in Carleton and Waksler (2000, 2002), non-pronominal objects are optionally preceded by the preposition *ji?i*. Carleton and Waksler (2002) argue in detail that the use of the preposition is determined by the information structure role of the object. They claim that *ji?i* marks an object as focused, perhaps on the basis of the (in our view, incorrect) assumption that a clause cannot contain more than one topic; however, they observe that their analysis necessitates a wholesale recasting of the notion of focus in order to account for the fact that the relevant referent is already present in the previous discourse. We believe that the *ji?i*-marked object is better accounted for by assuming that it is in fact topical and that what Carleton and Waksler refer to as "focus" is actually a topic according to our definitions.

Carleton and Waksler show that animacy of the object does not correlate with the distribution of *ji?i*: both animate and inanimate objects can be marked with the preposition, but in both cases it is optional. The correlation with specificity is also incomplete; although nonspecific objects do not seem to allow *ji?i*, specific objects are either marked or unmarked, as shown in (35):

- (35) a. Juan Ø-yu?u-nto:-yu (ji?i) Maria  
 Juan C-have-eye-3Sg Prep Maria  
 'Juan recognised Maria.' (Carleton and Waksler 2002:159)

- b. nkwi-ilo (ji?i) tonɛ?ɛ  
 C-rescue.3Sg Prep Loc.home  
 'He rescued the house.' (Carleton and Waksler 2002:159)

Examples (35a,b) also show that definiteness is not enough to ensure prepositional marking, since the objects in these examples are definite but not necessarily marked. Except for a few lexicalised expressions where objects are always nonspecific, the lexical semantics of the verb is usually irrelevant for object marking.<sup>2</sup>

The distribution described above is predicted if we assume that the postposition marks topical objects. As mentioned above, nonreferential (nonspecific) NPs are impossible as topics, and the postposition is impossible on such objects. As for specific objects, Carleton and Waksler (2002) further argue that the function of *ji?i* is to "zoom in one character", that is, to bring the object referent to the addressee's attention. They observe that the preposition signals to the addressee how to identify "a location in the discourse model where the speaker wants to bring the centre of the addressee's attention". This indicates that the object referent is pragmatically salient, as was suggested above for topical arguments. Moreover, it is highly likely to be present in the previous discourse.

In (36), the primary topic referent is the man who is the main character of the narrative:

- (36) nka-lo-yu ji?i na kuchilu-u? ntu-si?yu-yu yane  
 C-remove-Top Prep Det knife-Spec P-cut-Spec neck  
 'He took (his) knife and began to cut (his) throat.'  
 (Carleton and Waksler 2002:159)

This referent does not receive overt linguistic expression, but the verb bears the topic marker, indicating that its subject is the topic of the clause. The object 'knife' is marked as specific and is preceded by the definite article because it has been mentioned in the previous discourse. The narrative tells the story of an eagle and a serpent that terrorise the community. The second part of the narrative describes the plan to kill the eagle and the knife that was specifically purchased for this purpose. Thus, the pragmatic relation between the main character and the knife is well-established in the consciousness of the speaker

<sup>2</sup>Carleton and Waksler (2002:162) point out that the preposition tends not to occur on objects of verbs "whose lexical semantics entails most or all of the semantic features of the objects". These are verbs such as 'extinguish (fire)', 'earn (money)' or 'husk (corn)'. The objects of these verbs need not be individuated, since they are perceived as parts of the event denoted by the verb. As Carleton and Waksler (2002:163) note, "the speaker does not need to call the addressee's attention to that object, because its presence is already understood or expected".

and the addressee at the time (36) is produced. Carleton and Waksler state that, although the referent of ‘knife’ is not new, the relationship between this referent and the proposition is, which makes the knife “focused”. The first clause in (36) updates this relation: the clause is construed to be about the relation that holds between the man and his knife. By our criteria, the object is the secondary topic, and that is why it is marked by the preposition *ji?i*.

Note that in the second clause ‘and began to cut his throat’ the object ‘throat’ is also definite and specific, as indicated by the specific marker on the verb, but it has not been mentioned in the previous discourse and is not expected to be pragmatically salient at the time of the utterance. This object does not have topical status, but rather inhabits the domain of wide focus, and is not marked with the preposition.

Examples (37a) and (37b) are taken from the same text.

- (37) a. nte-su          nchi?yu nte-su          nchi?yu na nt-yotɨ-na  
 Hab-cut.3Sg fruit    Hab-cut.3Sg fruit    Neg Hab-know-1Pl  
 tukwi nchi?yu nte  
 what fruit    Hab.be  
 ‘He cuts fruit, he cuts fruit; we don’t know what kind of fruit it is.’
- b. nku-tyejna ntu-su-ka?a na nyatɛ-ɛ?    ji?i na nchi?yu  
 C-begin    Hab-cut-again Det person-Spec Prep Det fruit  
 ‘The man began to cut the fruit again.’
- (Carleton and Waksler 2002:167)

In (37a) the fruit is introduced for the first time, and the object preposition is not used. In (37b), which occurs later in the text, the relation between the subject referent and the object referent is already established and the object is pragmatically salient. This is the same fruit as was mentioned in (37a). The new information provided by (37b) is that cutting of the fruit by the man occurred again. This indicates that this fruit is under discussion when (37b) is produced and corresponds to the secondary topic.

Carleton and Waksler (2002) also discuss example (38), which appears in a context in which both the subject and the object are under discussion; in this example, the object receives topic marking.

- (38) ynɔ?ɔ ku-tze-ɔ?    ji?i kosa na tzo?o  
 hence P-fear-3Pl Prep thing Neg good  
 ‘Hence they fear (those) bad things/creatures.’
- (Carleton and Waksler 2002:166)

This example occurs in a narrative about hostile ‘semi-human’ creatures who prey on the women of the village, so the pragmatic relation between the villagers (the referent of the subject in (38)) and the creatures (the referent of the object) is well established. Sentence (38) occurs at the end of the narrative, after a number of other sentences that describe the relation between the villagers and the creatures. The function of the utterance in question is to update the addressee’s knowledge about this relation. In fact, Carleton and Waksler (2002:167) note that the preposition normally occurs “at the right boundary of a scene, sometimes, though not always, corresponding with the end of a narrative discourse”. This means that the referents involved in respective propositions are likely to be familiar to the interlocutors and the relationship between them is pragmatically established.<sup>3</sup>

Although marking on lexical objects depends on topicality, object personal pronouns must cliticise to the objective preposition, independently of their information structure status (Carleton and Waksler 2000). Thus, grammatical marking is extended to all pronominal objects in Chatino:

- (39) nk-yaja-nto-čʔ    jʔiʔi-chuʔ  
 C-found-eye-1Sg Prep-3SgFem  
 ‘I saw her (from afar).’                            (Carleton and Waksler 2000:390)

This is similar to the constraints in Mongolian requiring marking on all definite objects, independent of topicality; we discuss patterns such as these in Chapter 10.

We do not have much evidence about the syntactic properties of objects in Chatino. However, at least one piece of data suggests that casemarked and noncasemarked objects behave differently. Chatino is a verb-initial language, but Carleton and Waksler (2000) show that under certain pragmatic conditions, subjects as well as marked objects preceded by the preposition *jʔiʔi* can be dislocated by what they call “Focus Dislocation” into sentence-initial position. Recall that Carleton and Waksler call “focus” what we understand as topic, so it is likely that we are dealing with topicalisation rather than focus here. In (40), the subject appears sentence-initially:

- (40) tzaka ynatę    nte-su            nchiʔyu  
 one human Hab-cut.3Sg fruit  
 ‘A man is cutting fruit.’                            (Carleton and Waksler 2000:395)

---

<sup>3</sup>Carleton and Waksler (2002) also note that objects may be marked when disambiguation is needed between a complex subject consisting of two words, on the one hand, and the subject and the object, on the other hand. We do not address such cases here.

Marked objects (those marked with the preposition *jiʔi*) can also be fronted, while the preposition either appears together with the object or remains stranded.

- (41) a. *jiʔi kiʔyu nka-ra kunaʔa*  
 Prep man C-hit woman  
 ‘The woman hit the man.’
- b. *kiʔyu nka-ra kunaʔa jiʔi*  
 man C-hit woman Prep  
 ‘The woman hit the man.’ (Carleton and Waksler 2000:396)

On the other hand, objects unmarked with the preposition cannot appear in sentence-initial position. The position of the marked and unmarked object is otherwise the same, so this does not seem to be because unmarked objects must appear in a particular focus position. These patterns are similar to the Mongolian data discussed above, and seems to reflect a syntactic difference between the two kinds of objects: subjects and marked objects can be fronted, but unmarked objects cannot.

We suggest that the assignment of object grammatical functions depends on pronominality and topicality, and we assume the same treatment of the Chatino preposition *jiʔi* as for the Mongolian accusative affix - *ig*, given in (34) above: *-ig* marks its argument as OBJ.

### 8.3.3 *Hindi*

Hindi (Indo-Aryan) exhibits differential object casemarking coupled with a complicated agreement system (Mohanani 1994: Chapter 5). In example (42), the verb agrees with the uncase-marked subject:

- (42) a. *Ravii baalak-ko uʔhaaegaa*  
 Ravi.Masc boy-KO lift.Fut.3MascSg  
 ‘Ravi will lift up the boy.’
- b. *Niinaa baalak-ko uʔhaaegii*  
 Nina.Fem boy-KO lift.Fut.3FemSg  
 ‘Ravi will lift up the boy.’ (Mohanani 1994:103)

When the subject is case-marked, and there is a noncase-marked object, the verb agrees with the object:

- (43) *Ravii-ne roʔii khaayii*  
 Ravi-Erg bread.Fem eat.Perf.3FemSg  
 ‘Ravi ate bread.’ (Mohanani 1994:103)

When both the subject and the object are casemarked, the verb shows neutral (third person masculine singular) agreement:

- (44) Ravii-ne baalak-ko uṭhaayaa  
 Ravi-Erg boy-ko lift.Perf.3MascSg  
 ‘Ravi ate bread.’ (Mohanan 1994:103)

Mohanan (1994) demonstrates that agreement patterns in Hindi are dependent on casemarking patterns: on her analysis, the verb agrees with the uncase-marked argument which is highest on the thematic hierarchy. Casemarking on subjects of transitives and some intransitive verbs depends on aspect and agentivity. Our main concern in the following is the distribution of the case-marker *ko* on objects; *ko* is also possible on experiencer subjects and on subjects of passive verbs, but we will not be concerned with these uses of *ko* (see Mohanan 1994 for discussion of these patterns). We return below to a discussion of agreement patterns and how they can best be characterised.

Recall the basic pattern of object marking in Hindi, repeated here from Chapter 1.<sup>4</sup> As observed by Mohanan (1994), objects that bear *ko* are necessarily interpreted as specific. Human/animate specific objects are always marked, inanimate/nonhuman indefinite objects are always unmarked, and inanimate/nonhuman definite objects may be either marked or unmarked.<sup>5</sup>

- (45) Ilaa-ne bacce-ko/\*bacca uṭhaayaa  
 Ila-Erg child-ko/child lifted  
 ‘Ila lifted the/a child.’ (Mohanan 1994:80)

- (46) Ravii-ne kaccaa kelaa kaaṭaa  
 Ravi-Erg unripe banana cut  
 ‘Ravi cut the/a unripe banana.’ (Mohanan 1994:87)

- (47) Ravii-ne kaccaa kele-ko kaaṭa  
 Ravi-Erg unripe banana-ko cut  
 ‘Ravi cut the/\*a unripe banana.’ (Mohanan 1994:88)

Note that human/animate specific objects always take *ko* even if they are non-topical/focused, as shown in (48) and (49):

- (48) Hassan kis-ko/\*kaun maaregaa?  
 Hassan who-ko/who kill.Fut  
 ‘Whom will Hassan kill?’

<sup>4</sup>Examples for which the source is not given come from personal communication with Devyani Sharma and Tara Mohanan.

<sup>5</sup>There seems to be a certain amount of (dialectal) variation in this. For example, Butt (1993) states that object marking is possible for specific indefinite inanimates.

- (49) Hassan kisi-ko/\*koi maaregaa  
 Hassan someone-ko/someone kill.Fut  
 'Hassan will kill someone.'

For inanimate/nonhuman definite objects, casemarking appears to be optional; however, we argue that casemarking patterns are best explained by taking information structure roles into account. This fits with observations by McGregor (1972), who claims that object marking requires individualisation and a certain degree of contextual importance for the object, and with Butt and King's (1996) proposal that marking patterns are related to information structure roles. In their analysis, focused objects are unmarked and are licensed in the immediately preverbal position, which also hosts nonspecific objects. This is confirmed by our own data. Definite inanimate objects remain unmarked, or at least *ko* is strongly dispreferred, if the object is focused. The following sentences were produced as the answers to the question *What happened?*, and thus exemplify wide focus.

- (50) a. What happened?  
 b. Hassan-ne meraa kalam/?kalam-ko toḍ diyaa  
 Hassan-Erg my.Masc pen/pen-ko break gave  
 'Hassan broke my pen.'
- (51) a. What happened?  
 b. [jis kitaab ke-bare-mē aap baat kar rahe the] [voh  
 which book about-Loc you talk do Prog Past that  
 kitaab/?kitaab-ko mai-ne kharid li]  
 book/book-ko I-Erg buy take.Past  
 'I bought that book you were talking about.'

In the following examples, the object is in narrow focus.

- (52) a. What did Hassan sell?  
 b. Hassan-ne voh kitaab/?kitaab-ko bec diyaa  
 Hassan-Erg that book/book-ko sell gave  
 'Hassan sold that book.'
- c. Hassan-ne voh kursi/\*kursi-ko bec diyaa jo dukaan-mē  
 Hassan-Erg that chair/chair-ko sell give.Past Rel shop-Loc  
 thi  
 be.Past  
 'Hassan sold that chair that was in the shop.'



- (53) a. What is Anu doing?  
 b. Anu kitaab paḍh rahi hai  
 Anu book read Prog Pres  
 ‘Anu is reading the book.’ (Dayal 2003:84)

Again, there is no *ko*-marking on focused objects.

Specific non-focused objects can be either marked or unmarked. However, Butt and King (1996) say little about the conditions on their marking; they notice that accusative objects are generally more compatible with telic constructions and “affected” readings, but do not elaborate on this point. Our data show instead that aspectual differences are at best secondary, whereas contextual factors play a prominent role. In contexts where the object is salient and the utterance updates the addressee’s knowledge about the relation that holds between the subject and the object referents, *ko*-marking is required, as in example (52) above and the following example:

- (54) a. What did Hassan do to the book?  
 b. us kitaab-ko/us-ko/\*voh kitaab bec diyaa  
 that book-ko/that-ko/that book sell gave  
 ‘He sold it/that book.’

In (52) the object is in narrow focus and fills the informational gap in the presupposed open proposition ‘Hassan sold X’. In contrast, (54) is acceptable in the context of a certain pragmatically established relationship between the referent of ‘Hassan’ and the referent of ‘the book’. The utterance updates the information about this relationship and thus can be interpreted in our analysis as a secondary topic construction. There is no apparent aspectual difference between the two. The contrast between (52b) and (54b) indicates that objects of the same verb in the same tense/aspect can be either marked or unmarked, depending on their information structure role.

The examples below provide further illustration of the requirement for topical objects to be marked. Some of these examples additionally demonstrate that affectness is not directly relevant for marking: the object of the verb ‘see’ is the least affected by the event described in the sentences, but it behaves just like the highly affected object of the verb ‘break’ with respect to grammatical marking.

- (55) a. You should punish Hassan for breaking your pen.  
 b. mere kalam-ko/?kalam Ravi-ne toḍaa Hassan-ne nahī  
 my.Obl pen-ko/pen Ravi-Erg broke Hassan-Erg no  
 ‘It was Ravi who broke my pen, not Hassan.’

- (56) a. Did you read the book I was talking about?  
 b. abhi nahī, mai-ne us kitaab-ko/?kitaab kal hi kharidaa  
 now no I-Erg that book-ko/book yesterday Foc bought  
 ‘Not yet, I only bought that book yesterday.’
- (57) a. Did you buy that pen?  
 b. nahī, lekin mai-ne kalam-ko/us-ko/\*kalam dukaan-mē dekhaa  
 no but I-Erg pen-ko/that-ko/pen shop-Loc saw  
 ‘No, but I saw that pen/it in the shop.’

Consider also the following contrast:

- (58) a. ham mez pūch dēge  
 I.Hon table wipe give.Fut  
 ‘I am going to wipe the table.’
- b. mez-ko pūch do  
 table-ko wipe give.Imp  
 ‘Wipe the table!’

Both examples in (58) are acceptable in a situation in which the table has not been mentioned in the previous discourse but is present in the situation of speech and therefore identifiable and definite. However, the information structure role of the object differs. According to our consultants, the communicative purpose of the first utterance is to inform the addressee about the speaker’s intention. The object referent is construed as part of the event and is not individuated as a pragmatically salient element. Informationally, it is a part of wide focus, so *ko*-marking is absent. In the second case, the event is presented as a plan. The utterance is construed as an instruction to the addressee to perform a certain action with respect to the table, so the table plays a prominent pragmatic role in the situation. The object NP is topical and marked with *ko*. Thus, while indefinite nonhuman/nonanimate objects are unmarked, for definite inanimates *ko*-marking is not merely optional: it is determined by topicality of the object.

Marking with *ko* is also possible when the subject is in focus:

- (59) kalam-ko kis-ne toḍaa?  
 pen-ko who-Erg broke  
 ‘Who broke the pen?’

This example demonstrates that *ko*-marking is not necessarily confined to the secondary topic.

Butt and King were among the first to propose that monotransitive verbs can take either OBJ or OBJ<sub>θ</sub>, depending in part on information-structural factors. On their analysis, the factor determining the grammatical function of the object is position: an unmarked, nontopical object which is immediately preverbal has a different grammatical function from an object which may be marked, and whose position is unconstrained. They propose to analyse unmarked objects that are immediately preverbal as OBJ, and marked or unmarked objects whose position is unconstrained as OBJ<sub>θ</sub>.

Our analysis differs from Butt and King's in two respects. First, we believe that when marking correlates with a difference in grammatical function, the marked, topical object tends cross-linguistically to be associated with properties typical of core arguments, and the unmarked, nontopical object tends to exhibit fewer core properties: in the languages we have examined in this chapter, the marked, topical object is the primary object, OBJ, while the unmarked, focused object is a secondary object, OBJ<sub>θ</sub>. This is opposite to the distribution that Butt and King suggest for Hindi.

Second, we do not agree with Butt and King that the Hindi data motivate the postulation of a distinction in grammatical function based on phrase structure position, since we know of no independent evidence that would motivate such a distinction. Instead, we believe that Hindi is like the other languages examined in this chapter: the grammatical function of an argument is signalled by its case. Marked objects are OBJ, and unmarked objects are OBJ<sub>θ</sub>. This analysis makes possible a very simple statement of agreement patterns in Hindi. As shown in examples (42–44), the verb agrees with the subject unless it is case-marked; in that case, it agrees with the object unless it is casemarked; if both the subject and the object are casemarked, the verb shows neutral agreement. Mohanan (1994) analyses this pattern in terms of the position of the arguments of the verb on the thematic hierarchy: the verb agrees with the uncase-marked argument which is highest on the hierarchy. If we assume that marked objects are OBJ and unmarked objects are OBJ<sub>θ</sub>, we can restate this generalisation in purely syntactic terms: the verb agrees with the subject unless it is casemarked, in which case it agrees with OBJ<sub>θ</sub>, which is always unmarked.

The formal analysis of *ko* is exactly as in Mongolian and Chatino: *ko* marks the primary object, OBJ. Mapping rules for Hindi require animate and topical definite objects to be realised as OBJ, while inanimate and nontopical objects are OBJ<sub>θ</sub>.

#### 8.4 Objects and markedness

The discussion in this and previous chapters raises the question of functional markedness: which objects, topical (formally marked) or nontopical/focused (formally unmarked), count as prototypical, canonical, or functionally unmarked?

The prototypical alignment between syntax and information structure assumed in many functional and typological works is the SUBJ/TOPIC, OBJ/FOCUS alignment. In other words, it is commonly assumed that prototypical/canonical/functionally unmarked objects are in focus (Du Bois 1987, Bossong 1989, Lambrecht 1994, Maslova 2003b, Bossong 1991, Jelinek and Carnie 2003, and many others). Bossong (1991:158), for instance, argues that objects often “form an integrated part of the verbal complex”, so that the object and the verb belong to the same “pragmatic domain” (in our terminology, focus domain or comment). This idea is implicitly present in many analyses of DOM, in particular those that highlight its disambiguating function (see Chapter 1, Section 1.2.1), even though not many of them explicitly address information structure. For example, in Aissen’s (2003a, 2003b) analysis, prototypical (functionally unmarked) objects lack features that are ranked high on the Prominence Scales. But these features are known to characterise topics, so functionally unmarked objects are assumed to be nontopical/focused.

However, we have argued that objects tend to be more topical than other nonsubject grammatical functions. This idea obviously stands in contrast with the frequent claims that objects are prototypically in focus. Our analysis avoids this contradiction by relying on a more elaborated view of information structure than it is usually assumed and on the notion of secondary topic, as well as a richer inventory of objects, including primary objects (OBJ) as well as secondary objects (the family of grammatical functions OBJ<sub>θ</sub>).

We have shown that while subjects are normally primary topics, (primary) objects are the best candidates for secondary topichood. In Chapter 3 we discussed informational contexts in which secondary topic objects are likely to appear. They include argument-focus structures in which the focus is associated with an oblique element (e.g. *Where did he read the book?*) and structures in which focus is associated with the verbal element alone, to the exclusion of all nonverbal elements (the so-called “verum” focus, e.g. *He DID read the book*, and probably other structures as well. However, contexts that involve nontopical, focused objects are also frequent. They include utterances where the narrow focus falls on the object alone (answers to questions like *What did he write/read/eat/etc.?*) or broad focus structures where the object is included in the broad focus domain but is not pragmatically presupposed (*What did he do? He read a book*).

This functional “duality” has a syntactic corollary in some languages: this chapter has examined languages in which the distinction between marked and unmarked objects goes beyond grammatical marking and correlates with a difference in grammatical function. Topical, grammatically marked objects often pattern with subjects, and show more characteristics of core grammatical functions than nontopical objects, and we analyse the distinction in terms of the difference between the primary object OBJ and the secondary, restricted OBJ<sub>θ</sub>. In languages of this type, grammatically marked objects are primary objects: they are topical and can express various semantic roles. In contrast, unmarked objects are semantically restricted, nontopical (focused) secondary objects.

Assuming a grammatical function hierarchy with obliques toward the bottom and SUBJ/OBJ at the top, topics tend to link to arguments high on the hierarchy, while focus tends to go with lower ranked elements. Since objects are located in the middle of the hierarchy, they demonstrate a split: they are equally appropriate as topics or foci. This gives the following principles of default alignment between f-structure and information structure:<sup>6</sup>



In other words, the dual linkage of objects in the languages discussed in this chapter reflects a cross-linguistic tendency for splitting the object into two distinct grammatical functions depending on information structure role: a topical OBJ and a focused OBJ<sub>θ</sub>. This split is not universal, though: the languages discussed in Chapter 7 do not align information structure roles and grammatical functions, but rather display one informationally undifferentiated object function. In such languages semantic role takes priority over information structure role in mapping relations.

On this more articulated view, topical objects are not functionally marked; rather, we suggest that objects are equally likely to be topical and nontopical (focused). Thus, our analysis does not relate formal markedness on objects with their functional markedness. In fact, topical objects can be clearly shown to be functionally unmarked, if markedness is reflected in textual infrequency: topical objects are common in human discourse, and formally marked objects are just as frequent in languages with DOM as formally unmarked objects. This is supported by corpus data and corpus counts. In Givón’s (1979) text counts, 50% of objects were found to be definite; definiteness, of course,

<sup>6</sup>Börjars and Vincent (2008) propose to do away with the primary OBJ/secondary OBJ<sub>θ</sub> distinction, and treat all objects as instances of OBJ<sub>θ</sub>. Following this suggestion would not allow us to express the alignment generalisation in (60), which we believe to be an important fact about the information structure/syntax interface.

strongly correlates with topicality. More direct evidence comes from Tundra Yukaghir (isolate), as presented by Maslova (2003b). Yukaghir has a grammaticalised system of focus marking on subjects and objects in which the case marker of the subject or object indicates its focus vs. nonfocus role. This makes it particularly easy to identify the information structure role of an argument, based on its casemarking. Maslova presents the following frequency counts of information structure patterns in finite transitive clauses:

- (61) topical subject, topical object: ca. 65%  
 topical subject, focused object: ca. 35%  
 focused subject, topical object: less than 1%

These data demonstrate that topical objects (65%) are in fact **more frequent** than focused objects (35%), at least in the narrative texts on which the count was based.<sup>7</sup>

Roughly similar results have been obtained from Ostyak textual counts. The table below, updated from Nikolaeva (1999), shows object agreement patterns attested for transitive clauses in the collection of Ostyak texts recorded and published by Pápay (1906-1908).

| (62)   | no object agreement | object agreement |
|--------|---------------------|------------------|
| total  | 403 (40%)           | 611 (60%)        |
| (S)OV  | 329 (74%)           | 197 (26%)        |
| (S)OXV | 74 (15%)            | 414 (85%)        |

Agreeing/topical objects are slightly more frequent than nonagreeing/nontopical objects (60% in the first line of the table), so in this sense agreeing/topical objects cannot be considered functionally marked.

The table also demonstrates that in most (S)OV clauses — clauses containing an object and no other nonsubject elements — the object does not agree (74%); such objects can be analysed as focused. In the remaining 26% of (S)OV clauses the agreeing object is topical, while the focus domain includes the predicate alone, as in the English example (9c) discussed in Chapter 3 (page 55) and the Ostyak example (8) of this chapter. In (S)OXV clauses, the object normally agrees (85%); on our analysis, such objects are secondary topics. The element X is an oblique constituent, which is focused and normally located in the preverbal focus position.<sup>8</sup> This indicates that in clauses containing no other material than the subject, object and verb, the object is indeed a likely focus. Importantly, however, when an oblique constituent is present, it

<sup>7</sup>Maslova (2003b) states that although more than 97% of focus elements in transitive clauses are objects, most objects are encoded as topics.

<sup>8</sup>We also have a small number of (S)XOV examples, which require special explanation. Their number is insignificant, however.

is this constituent that tends to be associated with focus.<sup>9</sup> In such cases, the object has the status of secondary topic.

This suggests that “prototypical” alignment, which relates objects with the focus status, is mostly valid for clauses with no oblique constituents (SOV). When a clause contains one or several obliques, the focus function is typically associated with this oblique, while the object may be topical. This situation is predicted by our analysis and the notion of secondary topichood. We can see, then, that objects are **just as likely** to be topics as to be in focus.

## 8.5 Conclusion

Objects have two prototypical/canonical/unmarked information structure functions: focus and secondary topic. This, together with the fact that topics tend to be grammatically marked, provides an explanation for patterns of DOM in the languages we addressed in Chapter 7. Further, some languages restrict nonsubject topic marking to objects, and the realisation of an argument as a primary object (OBJ) or a secondary object (OBJ<sub>θ</sub>) can depend on its information structure role: in the languages examined in this chapter, topical arguments are realised as OBJ and nontopical arguments are realised as OBJ<sub>θ</sub>, insofar as this is consistent with additional constraints on argument realisation and expression and the inventory of grammatical functions in the language.

---

<sup>9</sup>In 15% of SOXV clauses, the object is in focus, while X is a component of a complex predicate and does not bear a distinct information structure role; see Nikolaeva (1999) for details.





## *Multiple objects and grammatical alignment*

Although the syntax of ditransitives and verbs with multiple complements is not the main focus of this book, we offer here some general observations on multiple object constructions in languages with DOM, and a review of alignment possibilities involving grammatical function, information structure role, semantic role, and marking for monotransitive and multitransitive constructions. For a thorough discussion of double object constructions in an LFG setting, see Börjars and Vincent (2008) and Lam (2008).

### **9.1 On the typology of multitransitive constructions**

In Chapter 2, Section 2.2, we reviewed the classic LFG theory of grammatical functions, which includes a distinction between primary and secondary objects, originally called OBJ and OBJ2 (Bresnan 1980, Kaplan and Bresnan 1982). The theory was subsequently refined by Bresnan and Kanerva (1989), who proposed to replace the OBJ/OBJ2 distinction with a distinction between the primary object OBJ and the family of secondary object functions OBJ<sub>θ</sub>.

The primary/secondary object distinction is also explored by Dryer (1986), who proposes a rough typology of argument linking patterns with trivalent verbs like ‘give’,<sup>1</sup> which take a theme and a goal. Dryer distinguishes two classes of ‘give’-type constructions. His first class draws the traditional distinction between a theme direct object and a goal indirect object; in LFG, this would be analysed as involving an OBJ theme and an oblique goal (OBL<sub>GOAL</sub>),

---

<sup>1</sup>Lam (2008) points out that in some languages, the verb ‘give’ is actually the least prototypical member of the class of trivalent verbs, rather than the most canonical member. We do not believe that this is true of the languages we discuss in this chapter.

as in (1a). This alignment pattern has also been referred to as “indirective”. Dryer’s second class distinguishes primary and secondary objects (OBJ vs. OBJ<sub>THEME</sub>), as in (1b). This pattern is commonly referred to as “ditransitive”, and also been termed “secundative” (Haspelmath 2007, Malchukov et al. 2007; see also Lam 2008).

- (1) a. I gave a book to Bill. (indirective)  
 b. I gave Bill a book. (secundative)

These can be schematically represented as in (2):

- (2)
- |                                |                   |                                    |
|--------------------------------|-------------------|------------------------------------|
| indirective (Dryer’s class 1): | THEME<br> <br>OBJ | GOAL<br> <br>OBL-GOAL              |
| secundative (Dryer’s class 2): | GOAL<br> <br>OBJ  | THEME<br> <br>OBJ <sub>THEME</sub> |

Dryer’s analysis of English examples like (1b) has been challenged by Börjars and Vincent (2008), who argue (following work by Hudson 1992) that the theme in examples like (1b) behaves more like a primary object than the goal does. The behavioural evidence for the grammatical status of object-like arguments which we have examined in Chapters 7 and 8 underpins our discussion of alignment patterns for trivalent ‘give’-type verbs in the following, though more work is clearly needed on the syntactic status and classification of objects in multitransitive constructions, even for the English construction exemplified in (1b).

Kittilä (2006a) elaborates on Dryer’s typology, exploring casemarking patterns with recipients, goals, patients, and themes and how they are affected by the “animacy” of these arguments; for him, animacy is a cover term which also encompasses definiteness and a general notion of prominence. He identifies three casemarking strategies for the complements of monotransitive and ‘give’-type trivalent verbs. In the “role-based” strategy, the theme of a ‘give’-type verb and the patient of a monotransitive verb are marked in the same way, and the goal is marked differently. In the second type, which Kittilä calls “object-based”, the patient argument of a monotransitive verb and both the goal and the theme of a ‘give’-type trivalent verb bear object marking, independent of their animacy; this strategy requires all object-like arguments to be marked in the same way. In the third type, called “animacy-based”, animate (definite, prominent) patients, themes and goals are marked, while inanimate patients, themes and goals are unmarked. For monotransitives, this gives rise to DOM. There are two subtypes of this strategy for trivalent verbs, under the

assumption that the goal is always animate and marked: in “extended differential object marking”, both the goal and the animate theme are marked, while the inanimate theme remains unmarked; in “shifted differential object marking”, the goal is marked, but the theme is unmarked, even if it is animate. In a related paper, Kittilä (2006b) explores animacy-based marking patterns in examples where both the goal and theme of a ‘give’-type verb are animate. For such examples, he distinguishes “theme-prominent” languages, those in which the theme is an object and the goal is an oblique, from “recipient-prominent” languages, where the recipient is an object and the theme bears a different grammatical function.

Kittilä’s classifications are primarily based on morphological marking (case and/or agreement) associated with core arguments, which is taken to be definitional as a diagnostic of grammatical function for ‘give’-type verbs. The same is true of Haspelmath (2007) and Malchukov et al. (2007), who discuss patterns of ditransitive alignment from the point of view of what they refer to as “coding” (case or adpositional marking, agreement and position). From a marking/coding standpoint, there is also room for an additional “neutral” alignment, in which the monotransitive theme, the ditransitive theme and the goal all pattern in the same way.

However, the picture is more complicated if we recognise that grammatical marking and grammatical functions are independent parameters: LFG does not assume that identical casemarking and agreement patterns entail identical grammatical properties (Mohan 1982, Andrews 1982, Zaenen et al. 1985). Work within LFG on English and many other languages shows clearly that grammatical marking, including casemarking, agreement, and phrasal position, does not necessarily correspond one-to-one with grammatical function. Indeed, in our analysis, “theme-prominent” languages may be divided into two subtypes.<sup>2</sup> In the first type, marking differences do not correlate with a difference in grammatical function, and the marked and the unmarked theme both correspond to the primary object, as discussed in Chapter 7. In the second type, the marked theme is OBJ and the unmarked theme is OBJ<sub>θ</sub>, as discussed in Chapter 8. Kittilä and many other authors do not explore the possibility that a difference in marking for the object of a monotransitive verb might correlate with a difference in grammatical function, although we have shown that in some languages DOM itself can signal a difference in grammatical function. More generally, existing typologies do not extend to languages like those examined in Chapter 8, in which a monotransitive verb can take either a primary OBJ or a secondary OBJ<sub>θ</sub> (see also Çetinoğlu and Butt 2008 and Dahlstrom 2009).

---

<sup>2</sup>We have not observed a similar split in “recipient-prominent” languages: our data contains no clear examples in which the ditransitive goal exhibits DOM.

Our goal in the following is not to present a typology of ditransitive constructions in languages with DOM, but to show that it is likely to be more elaborate than has been previously thought. A complete typology of ditransitive constructions must take several factors into account: the inventory of grammatical functions available in the language, the mapping rules relating grammatical functions in f-structure to semantic roles at argument structure, the relationship between grammatical functions and grammatical marking, and the availability of double object constructions.

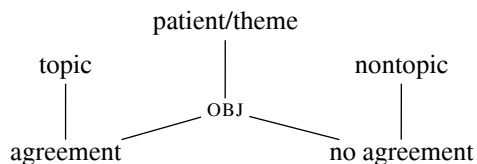
## 9.2 No ditransitive construction

Several of the languages we have examined in Chapters 7 and 8 lack a ditransitive/secundative construction, and make use only of indirect alignment for ‘give’-type verbs, involving a theme object and a goal oblique argument. In Nenets there is only one object function, and a ditransitive construction is therefore not available. Surprisingly, there are languages which have two object functions, and in which monotransitive verbs can take either a primary OBJ or a secondary OBJ<sub>θ</sub>, but which nevertheless do not have a double object construction: in these languages, a verb cannot take a primary object and a secondary object at the same time. Ostyak and Mongolian exemplify this pattern. These languages show that the availability of a double object construction in a language may depend on factors beyond the inventory of grammatical functions in the language.

### 9.2.1 Nenets

As shown in Chapter 7, Section 7.2, marked and unmarked monotransitive objects in Tundra Nenets correspond to the same grammatical function. Nenets has only one object function, LFG’s OBJ. Alignment involving semantic role, information structure role, grammatical function, and grammatical marking for Nenets monotransitives with third person objects can be graphically represented as in (3); recall that the verb never agrees with first or second person objects.

(3) Nenets monotransitives with third person objects:



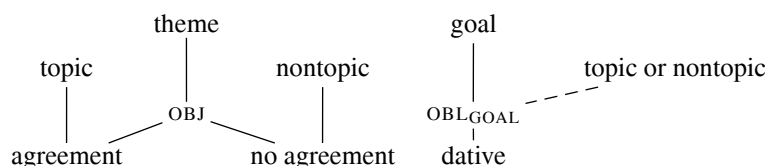
A third person object argument maps to OBJ, whether or not it is topical.<sup>3</sup> The presence or absence of agreement depends on information structure role: the verb agrees with topical objects, and not nontopical objects. Agreement patterns are, then, determined by a combination of person, grammatical function, and information structure.

The goal argument of ‘give’-type verbs is expressed by a dative-marked oblique:

- (4) Petya Masha-n<sup>h</sup> ti-m            myiqqa        / myiqqada  
 Peter Masha-Dat reindeer-Acc give.3SgSubj    give.Obj.3SgSubj  
 ‘Peter gave Masha a/the reindeer.’

The dative-marked oblique goal, OBL<sub>GOAL</sub>, displays no object properties: it cannot be passivised or relativised by means of the primary strategy relativising objects, and it does not participate in object control constructions. The alignment pattern for ‘give’-type verbs is shown in (5): topical/agreeing and nontopical/nonagreeing objects are OBJ, while the goal argument is an oblique goal OBL<sub>GOAL</sub>, independent of its information structure role:

- (5) Nenets ‘give’-type verbs with third person objects:



As this diagram shows, information structure does not affect the syntactic realisation of the goal argument: it is an oblique goal, whether or not it is topical. Further, there is no effect of the presence of the goal argument on agreement possibilities for the OBJ: exactly as in the monotransitive construction, the verb agrees with a topical OBJ and not a nontopical OBJ.

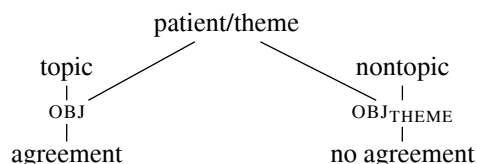
### 9.2.2 *Ostyak*

In Chapter 8, Section 8.2, we saw that agreeing and nonagreeing monotransitive objects in Northern Ostyak realise different grammatical functions: agreeing objects are primary objects, OBJ, while nonagreeing theme objects are secondary objects, OBJ<sub>THEME</sub>. Agreement depends purely on grammatical function. Information structure role determines whether theme arguments of mono-

<sup>3</sup>As far as we know, there are no verbs in Nenets which take an object with a thematic role other than patient/theme.

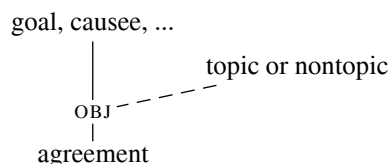
transitive verbs are realised as OBJ or OBJ<sub>THEME</sub>: topical themes are OBJ, while nontopical themes are OBJ<sub>THEME</sub>.

(6) Ostyak monotransitives with patient/theme objects:



We assume that Ostyak has only one restricted object, OBJ<sub>THEME</sub>, and no other objective functions: no OBJ<sub>GOAL</sub> or OBJ<sub>CAUSEE</sub>, for example. Since Ostyak has only OBJ<sub>THEME</sub>, objects with other semantic roles — for example, causees — must be realised as agreeing OBJ, since the grammatical role of OBJ<sub>THEME</sub> is reserved for themes. Alignment for monotransitives in Ostyak is as follows:

(7) Ostyak monotransitives with causee or goal objects:



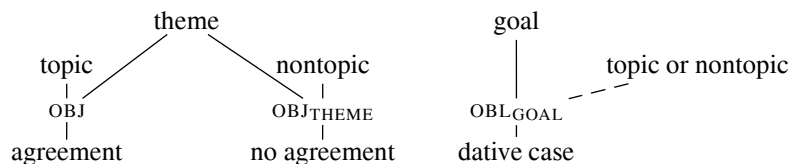
Even though Ostyak has two object functions, they cannot appear together, as arguments of the same verb: Ostyak does not have a double object construction. With verbs like ‘give’, either the goal or the theme must appear as an oblique, resulting in two possibilities. If the theme is an object, the goal is a dative oblique, as in (8); in that case, the goal can be either topical or nontopical.

(8) [What did you do to the cup? or What did you do to Peter?]

ma a:n Pe:tra e:lti ma-s-e:m / ma-s-əm  
 I cup Peter to give-Past-Obj.1SgSubj give-Past-1SgSubj

‘I gave a/the cup to Peter.’

(9) Ostyak ‘give’-type verbs with theme object:



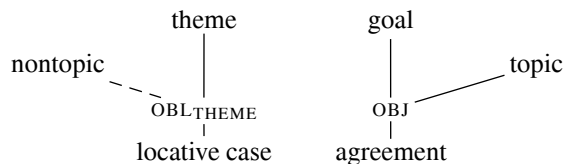
If the goal is realised as an object, it must be topical, and must agree; in that case, the theme is a locative oblique, as in (10). Recall that the subject in Ostyak is closely associated with the primary topic; hence, when the goal is topical, the theme is associated with a nontopic role (assuming that utterances with three topical arguments — subject, object, and oblique — are very rare or nonexistent).

(10) [What did you do to Peter?]

ma Pe:tra a:n-na ma-s-e:m / \*ma-s-əm  
 I Peter cup-Loc give-Past-Obj.1SgSubj give-Past-1SgSubj

‘I gave Peter a/the cup.’

(11) Ostyak ‘give’-type verbs with goal object:



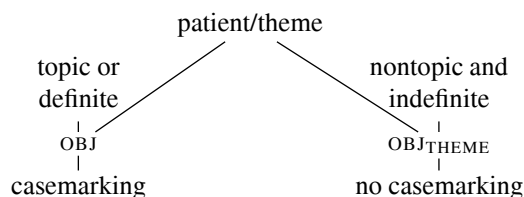
It is clear that agreement patterns depend entirely on grammatical function: the verb agrees with OBJ, but not with OBJ\_THEME or obliques. As in Nenets, the presence of an oblique goal argument does not affect agreement possibilities for object themes: the verb agrees with the topical OBJ, but not the nontopical OBJ\_THEME. When the object is a goal, as in (10), or a causee, as in examples (19) and (20) in Chapter 8 (page 154), it is a primary object, OBJ, and must agree with the verb. In this case, the theme cannot appear as OBJ, since that role is filled by the goal and cannot be doubly filled. Ostyak differs from Nenets in allowing the theme to be expressed as an oblique, OBL\_THEME; this allows the pattern diagrammed in (11), where the goal is an agreeing primary OBJ, and the theme is an oblique. The existence of this possibility means that the relation between information structure role and grammatical function is obscured even

for a verb like ‘give’, since the primary OBJ in the ‘give’-construction can be either a topical theme, or a goal with unspecified information structure role.

### 9.2.3 *Mongolian*

As shown in Chapter 8, Section 8.3.1, Mongolian monotransitive theme objects are primary objects (OBJ) and casemarked if topical or definite, and secondary objects (OBJ<sub>THEME</sub>) and noncasemarked if nontopical and indefinite. Diagrammatically, alignment between different grammatical levels in Mongolian monotransitives is similar to Ostyak:

(12) Mongolian monotransitives:



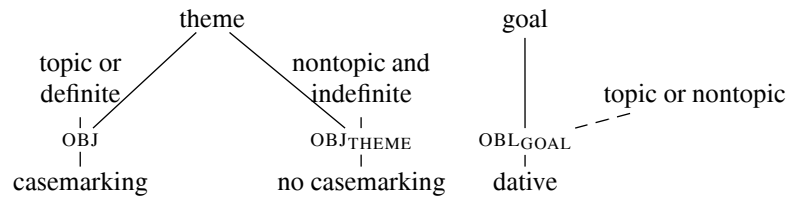
Like Ostyak, Mongolian does not have a ditransitive construction; again, we see that languages can have both primary and secondary objects in their inventory of grammatical functions without also having a ditransitive construction, where the two kinds of objects appear together. The goal argument of a three-argument verb such as ‘give’ in Mongolian must be expressed by an oblique, marked with the dative *-d*, and not an object:

(13) Bi Toya-d nom(-ig) og-son  
 I Toya-Dat book(-Acc) give-Past  
 ‘I gave Toya a/the book.’

The possibilities for expression of the arguments of the verb ‘give’ in Mongolian are similar to Nenets, and more restricted than in Ostyak. The primary object in Ostyak can correspond to the goal, with the theme realised as a locative oblique. In Mongolian the goal cannot be realised as the object of a trivalent verb such as ‘give’, because Mongolian does not have a ditransitive construction and also does not allow oblique themes. The goal of Mongolian ‘give’-type verbs must be marked by dative case, as in Nenets, and we analyse the dative-marked argument as an oblique goal, OBL<sub>GOAL</sub>.



(14) Mongolian ‘give’-type verbs:



Since the goal argument is realised as an oblique phrase in this construction, there is no obstacle preventing the theme of a ‘give’-type verb from surfacing as either a primary OBJ or a secondary OBJ<sub>THEME</sub>: it may be either marked (OBJ) or unmarked (OBJ<sub>THEME</sub>), depending on definiteness and information structure role.

All three languages examined in this section, Ostyak, Nenets and Mongolian, can be classified as theme-prominent languages from Kittilä’s (2006b) perspective. However, only in Nenets does the theme invariably correspond to the primary object. In Ostyak and Mongolian, grammatical functions are closely aligned with casemarking, and the theme is either marked OBJ or unmarked OBJ<sub>THEME</sub>.

### 9.3 Ditransitive constructions: Goal as primary object, theme as secondary object

Although LFG does not treat casemarking as an unambiguous indicator of grammatical function in every language, casemarking (together with other factors) is often a good indication of grammatical function, as demonstrated by casemarking patterns in Ostyak and Mongolian. Similarly, object casemarking in Chatino and Hindi correlates with grammatical function, as shown in Chapter 8: the marked object of a monotransitive verb is the primary OBJ, and the unmarked object is the secondary OBJ<sub>θ</sub>. In languages which have a ditransitive construction and also exhibit a tight alignment between marking and grammatical function, we do not expect to find both objects of a ditransitive verb marked in the same way, since it is not possible for both objects to correspond to the same grammatical function. Further, we expect the primary object of a ditransitive verb to be marked in the same way as the primary object of a monotransitive verb. This is what we find in Chatino and Hindi: in both languages, the goal of a ‘give’-type verb is marked in the same way as the OBJ of a monotransitive verb, while the theme remains unmarked. This alignment corresponds to the recipient-prominent type in Kittilä’s (2006b) typology. In fact, it is cross-linguistically common for goals and topical monotransitive objects

to be marked similarly, and Section 9.5 of this chapter discusses the pragmatic and historical basis for this similarity.

Alignment patterns for Chatino and Hindi are, then, similar in some respects to ‘give’-type verbs in Ostyak with goal objects, diagrammed in (11) of this chapter. The crucial factor distinguishing Ostyak from Chatino and Hindi is that while Ostyak does not have a double object construction, Chatino and Hindi do.

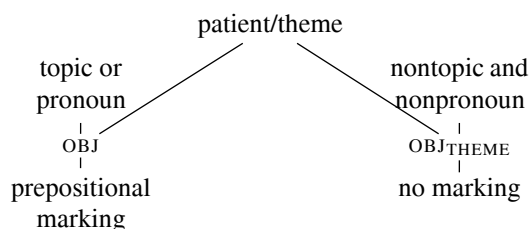
### 9.3.1 *Chatino*

In Chapter 8, Section 8.3.2, we saw that the object of a monotransitive verb in Chatino is marked with the preposition *jiʔi* if it is either pronominal or topical, and we argued on the basis of the “Focus Dislocation” construction that marked and unmarked objects correspond to different grammatical functions, OBJ and OBJ<sub>θ</sub> respectively.

- (15) Juan  $\emptyset$ -yuʔu-nto:-yu (jiʔi) Maria  
 Juan C-have-eye-3Sg Prep Maria  
 ‘Juan recognised Maria.’ (Carleton and Waksler 2002:159)

This can be diagrammatically represented as in (16):

- (16) Chatino monotransitives:



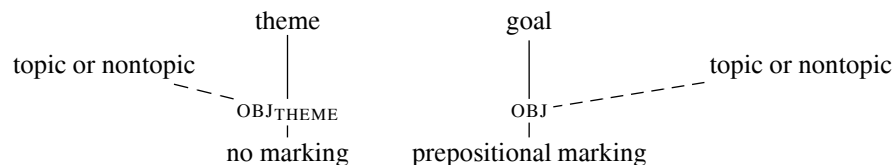
The same preposition marks not only objects of monotransitive verbs, but also goal/beneficiary objects of verbs such as ‘give’ or ‘write’. Here, the preposition is obligatory:

- (17) nka-ta na kiʔyu jne jiʔi/\* $\emptyset$  Juʔ  
 C-give.3Sg Det man money Prep Juan  
 ‘The man gave the money to Juan.’ (Carleton and Waksler 2000:388)

- (18) ŋka-sʔʔʔ kiti iʔi/\* $\emptyset$  Juʔ  
 C-write.3Sg letter Prep Juan  
 ‘She wrote a letter to Juan.’ (Carleton and Waksler 2000:389)

Carleton and Waksler do not explicitly state that prepositional marking of the theme in these examples is impossible, but they do not provide any examples in which both arguments are marked. Under the assumption that only the goal can be marked with preposition *ji?i* and the theme must remain unmarked, we can analyse the goal as the primary OBJ, in keeping with the tight correspondence in monotransitive constructions between *ji?i*-marking and the OBJ function. We treat the unmarked theme as the OBJ<sub>THEME</sub>, just as in monotransitive constructions.

(19) Chatino ‘give’-type verbs:



### 9.3.2 Hindi

Hindi, like Chatino, marks primary and secondary objects differently in monotransitive constructions: the casemarked object of a monotransitive verb in Hindi is the primary OBJ, and the uncasemarked object is the secondary OBJ<sub>θ</sub>, as discussed in Chapter 8, Section 8.3.3.

- (20) a. Ravi-ne kaccaa kelaa kaṭaa  
 Ravi-Erg unripe banana cut  
 ‘Ravi cut the/a unripe banana.’ (Mohanani 1994:87)
- b. Ravi-ne kaccaa kele-ko kaṭaa  
 Ravi-Erg unripe banana-KO cut  
 ‘Ravi cut the/\*a unripe banana.’ (Mohanani 1994:88)

In standard (prescriptive) Hindi (Masica 1982:20, Kellogg 1955 [1893]:399), ditransitive constructions cannot contain two *ko*-marked objects: the goal must be marked and the theme must be unmarked. As in Chatino, this is exactly what we expect if the two objects correspond to different grammatical functions, signalled by different casemarking patterns.

- (21) a. ilaa-ne maã-ko yah haar diyaa  
 Ilaa-erg mother-KO this necklace gave  
 ‘Ila gave this necklace to mother.’ (Mohanani 1994:85)

- b. \*ilaa-ne maā-ko is haar-ko diyaa  
 Ilaa-erg mother-KO this necklace-KO gave  
 ‘Ila gave this necklace to mother.’ (Mohanani 1994:85)

Assuming that casemarking is a reliable indicator of grammatical function in Hindi, we analyse Hindi in the same way as Chatino: the marked goal is OBJ, and the theme argument is OBJ<sub>THEME</sub> (exactly as shown in 19 for Chatino).

However, at least some nonstandard varieties of Hindi do allow the marking of both objects by the same element *ko*, as shown in example (22):

- (22) admii-ne us kitaab-ko aurat-ko diyaa  
 man-Erg that.Obl book-KO woman-KO give.Past.MSg  
 ‘The man gave that book to the woman.’ (Kittilä 2006b:302)

Similar data are cited by Bhatt and Anagnostopoulou (1996), who argue that, despite identical marking, the theme and the goal argument have different properties: for instance, only the theme object can be moved to the pre-subject position via syntactic topicalisation.

- (23) Tim-ko Theo-ne Sita-ko diyaa  
 Tim-KO Theo-Erg Sita-KO give.Past.MSg  
 ‘Tim, Theo gave to Sita.’  
 NOT: \*‘Sita, Theo gave to Tim.’  
 (Bhatt and Anagnostopoulou 1996:(19))

In these varieties of Hindi, *ko*-marking cannot be analysed as unambiguously marking arguments as OBJ; instead, marking depends on a combination of syntactic and information structure role, as in the languages discussed in Chapter 7. In Kittilä’s classification, standard Hindi belongs to the recipient-prominent type (goal is the primary OBJ). Nonstandard Hindi is likely to be a theme-prominent language (theme is the primary OBJ, and goal bears some other grammatical function), with no one-to-one alignment between grammatical function and casemarking, similar to Nenets. Further work is needed to explore alignment patterns and the syntactic behaviour of marked and unmarked objects in these nonstandard varieties, but it is possible that, unlike Nenets, nonstandard Hindi has a double-object construction with the theme as primary object and goal as a restricted object (possibly OBJ<sub>GOAL</sub>).

#### 9.4 Other multitransitive constructions

Alignment patterns with ‘give’-type verbs in the languages we have examined in the previous sections are, for the most part, easily classified in terms of

the two types of constructions examined by Dryer (1986), and echo the familiar patterns of the English dative alternation exemplified in (1). In Nenets, Mongolian, and the theme=object alignment pattern in Ostyak, examined in Section 9.2, the goal of a 'give'-type verb corresponds to an oblique phrase, and the theme behaves in the same way as a monotransitive object, with its casemarking and, in Ostyak and Mongolian, its grammatical function (OBJ or OBJ<sub>THEME</sub>) dependent on topicality. This is the object/oblique or direct object/indirect object alignment, shown for English in example (1a) of this chapter. In contrast, the languages examined in Section 9.3 (Chatino and Hindi) employ the primary object/secondary object, ditransitive alignment, with the goal as the primary object and the theme as the secondary object, shown for English in example (1b) of this chapter. In Chatino and Hindi, grammatical function is tightly aligned with marking: the goal is marked in the same way as the monotransitive OBJ, and the theme is marked in the same way as the monotransitive OBJ<sub>THEME</sub>. We now turn to a discussion of languages that differ in interesting ways from these familiar patterns, though all of them allow constructions with multiple objects.

We showed in Chapter 7 that marking in Dolakha Newar and Tigre for monotransitive verbs does not correlate with grammatical function, but only with information structure role: marking depends on topicality, but marked and unmarked objects both correspond to the same grammatical function, OBJ. Since there is a looser connection between casemarking and grammatical function for objects in these languages, it is not surprising that the theme and the goal in a 'give'-type construction can both be marked in these languages, as in the nonstandard varieties of Hindi discussed in the previous section. In both languages, the theme and the goal of a 'give'-type verb both exhibit a range of properties associated with objects, and it proves difficult to make a conclusive identification of the object-like grammatical function borne by goal and theme objects. We provide some discussion of object tests and the classification of objects of 'give'-type verbs for both languages.

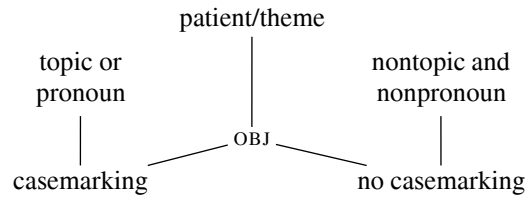
Many languages have an applicative construction, briefly discussed in Chapter 7, Section 7.1, where a nonobject argument is promoted to object via applicativisation; the resulting construction often has more than one object. In some languages, applicativisation is obligatory when a nonsubject argument is topical. In Section 9.4.3, we discuss object marking and topicality in one such language, Upper Nexaca Totonac (Totonac-Tepehua).

#### 9.4.1 Dolakha Newar

Chapter 7, Section 7.3.2 discussed casemarking patterns in Dolakha Newar, and concluded that (as in Nenets) there is no evidence for a difference in grammatical function between marked and unmarked theme objects in monotransi-

tive verbs: both are OBJ, with casemarking patterns depending on information structure role and pronominality (pronouns are obligatorily marked).

(24) Dolakha Newar monotransitives with nonpronominal objects:



With ‘give’-type verbs in Dolakha Newar, as in Chatino and Hindi, grammatical marking of the goal argument is morphologically identical to the marking on the monotransitive topical patient/theme. Unlike the monotransitive object, the goal must be marked, no matter what its pragmatic or semantic characteristics; omission of the object marker is reported to be impossible. This is true not only for objects of ‘give’-type verbs, but also for objects of exchange and beneficiary arguments:

- (25) a. *yā-ta dyābā bi-en ta-u-ī*  
 rice-Obj money give-Part put-Past.1Sg  
 ‘I had given money for the rice.’ (Genetti 1994:52)
- b. *ināgu khā-ri guntaṅ da-hat*  
 this.type matter-Ind nobody.Obj Proh-say  
 ‘Don’t tell anyone about this type of matter.’ (Genetti 2007:114)

Given the lack of a tight correlation between grammatical function and marking in monotransitives, we might expect identical morphological marking on the goal and the theme of a ditransitive clause, if appropriate informational conditions are met. This is exactly what is observed: object marking can occur twice within the same ditransitive clause, as shown in (26):

- (26) *āle āmta bhānche-ta bir-ju*  
 then 3Sg.Obj cook-Obj give-Past.3Sg  
 ‘Then he gave her (in marriage) to the cook.’ (Genetti 2007:316)

According to Genetti, patient and goal objects in ditransitive constructions share many syntactic properties, and she analyses them as instantiating the same grammatical function: object. On this view, Dolakha Newari represents the so-called neutral alignment, where all three object-like arguments pattern identically. This conclusion is based on the following three considerations.

This conclusion is based on the following three considerations. First, the two types of object can bear identical casemarking, as evidenced by example (26), though casemarking on non-patient/non-theme objects is obligatory and not optional.

Second, both types of object can antecede the emphatic reflexive element *āme tuḡ* ‘his/her own’.

- (27) a. *rām-na mucā āme tuḡ mā-ta bir-ju*  
 Ram-Erg child 3Sg.Gen Foc mother-Obj give-Past.3Sg  
 ‘Ram gave the child to his own (child’s) mother.’  
 (Genetti 1994:317)
- b. *rām-na krisna-ta āme tuḡ kitāb bir-ju*  
 Ram-Erg Krishna-Obj 3Sg.Gen Foc book give-Past.3Sg  
 ‘Ram gave Krishna his own (Krishna’s) book.’ (Genetti 1994:317)

In (27a) the reflexive is controlled by the unmarked theme object, and in (27b) by the marked goal object. Third, both types of object are relativised using the same relativisation strategy, and are the only arguments to be relativised in that way.

We draw a different conclusion from the results of these tests: we believe that the objecthood tests discussed by Genetti pick out the entire class of object functions, primary OBJ as well as secondary OBJ<sub>θ</sub>. In other words, Genetti’s tests pick out those arguments which are “objective” (+o) in LFG terms: see Chapter 2, Section 2.2. In an LFG setting, Genetti’s claim that the patient and the goal object of ‘give’-type verbs correspond to exactly the same grammatical function is untenable; two different semantic roles cannot correspond to the same grammatical function. For theory-internal reasons, then, we must analyse the goal and the theme as corresponding to two different object-like grammatical functions. In fact, there is at least one behavioural difference between the two: marking on the goal argument is obligatory and independent of information structure role, while marking on the theme argument is optional and correlates with topicality. We believe that this provides empirical motivation for the claim that the theme and the goal correspond to different object-like functions. The next question is what those object-like functions are.

The first possibility is that the object of a monotransitive verb and the theme object of a ‘give’-type verb correspond to OBJ<sub>THEME</sub>, and the goal argument corresponds to OBJ (this does not match the analysis which we presented in Chapter 7, Section 7.3.2). On this analysis, casemarking is obligatory for OBJ, and correlates with topicality for OBJ<sub>THEME</sub>.

(28) Analysis 1 (to be rejected), Dolakha Newar:

|                                    |                                                         |
|------------------------------------|---------------------------------------------------------|
| Monotransitive:                    | Ditransitive:                                           |
| THEME<br> <br>OBJ <sub>THEME</sub> | THEME    GOAL<br>          <br>OBJ <sub>THEME</sub> OBJ |

We discount this possibility on two grounds: it is implausible that the theme argument would never correspond to the primary OBJ (see Börjars and Vincent 2008 for discussion of the close relation between OBJ and theme), and it is also implausible that primary OBJ would appear only in a ditransitive construction, and not the monotransitive construction.

The second possibility is that the object of a monotransitive verb and the theme object of a ‘give’-type verb are OBJ, and the goal argument is OBJ<sub>GOAL</sub>. This would make ‘give’-type constructions in Dolakha Newar the reverse of Hindi or Chatino:

(29) Analysis 2:

|                                    |                   |                                    |                                  |
|------------------------------------|-------------------|------------------------------------|----------------------------------|
|                                    | Monotransitive:   | Ditransitive:                      |                                  |
| Dolakha Newar:                     | THEME<br> <br>OBJ | THEME<br> <br>OBJ                  | GOAL<br> <br>OBJ <sub>GOAL</sub> |
|                                    | Monotransitive:   | Ditransitive:                      |                                  |
| For comparison,<br>Hindi, Chatino: | THEME<br> <br>OBJ | THEME<br> <br>OBJ <sub>THEME</sub> | GOAL<br> <br>OBJ                 |

On Analysis 2, marking patterns are easy to account for: marking on OBJ depends on topicality, while marking on OBJ<sub>GOAL</sub> is obligatory. This is the analysis we hypothesized in Section 9.3.2 for non-standard Hindi.

The third possibility is that alignment in Dolakha Newar is exactly like Hindi and Chatino: the goal is OBJ, and the theme is OBJ<sub>THEME</sub>:

(30) Analysis 3, Dolakha Newar:

|                   |                                                         |
|-------------------|---------------------------------------------------------|
| Monotransitive:   | Ditransitive:                                           |
| THEME<br> <br>OBJ | THEME    GOAL<br>          <br>OBJ <sub>THEME</sub> OBJ |

This analysis has the virtue of making Dolakha Newar conform to alignment patterns attested in other languages: not only Hindi and Chatino, but also English, at least on the secundative analysis advocated in traditional LFG analyses (Bresnan 1980, Kaplan and Bresnan 1982) and by Dryer (1986). The cost is

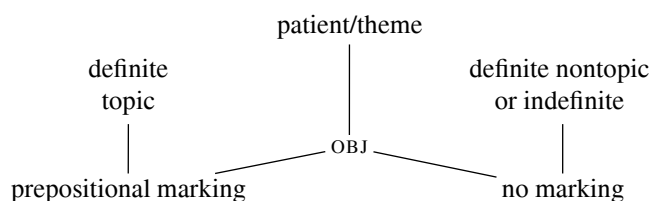


a considerable complication of the constraints on marking: on this analysis, *-ta*-marking is obligatory for primary objects which are goals, and dependent on topicality for theme objects, whether OBJ or OBJ<sub>THEME</sub>. Without additional data, we are unable to determine whether Analysis 2 or Analysis 3 better reflects the behaviour of objects in ‘give’-type constructions in Dolakha Newar.

#### 9.4.2 Tigre

As in Dolakha Newar, marking of monotransitive objects in Tigre depends not on grammatical function, but on topicality and definiteness. We showed in Chapter 7, Section 7.3.1 that prepositionally marked objects are definite and topical, and unmarked objects are indefinite or definite nontopical:

(31) Tigre monotransitives:



The Tigre verb shows agreement with definite objects, but this does not depend on topicality: definite objects control agreement, while indefinite objects do not. Interestingly, with ‘give’-type verbs the controller of agreement is the leftmost object in terms of its linear position, independent of its semantic role as theme or goal. In both (32a) and (32b) the leftmost argument controls agreement: the goal in (32a), and the theme in (32b):

- (32) a. Lilat    ?igil °ali    la waraqat habetto    /  
 Lilet.Fem Prep Ali.Masc the letter.Fem gave.3Fem.3Masc  
 \*habettā  
 gave.3Fem.3Fem  
 ‘Lilat gave Ali the letter.’ (Jake 1980:75)
- b. Lilat    la waraqat ?igil °ali    habettā    /  
 Lilet.Fem the letter.Fem Prep Ali.Masc gave.3Fem.3Fem  
 \*habetto  
 gave.3Fem.3Masc  
 ‘Lilat gave the letter to Ali.’ (Jake 1980:75)



- (36) Lilat la sa<sup>c</sup>at ?igil la ?inās habetta<sup>-</sup>  
 Lilet.Fem the watch.Fem Prep the man.Masc gave.3Fem.3Fem  
 (\*tu)  
 (\*3Masc)  
 ‘Lilat gave the watch to the man.’ (Jake 1980:78)

Thus, the theme object of a ‘give’-type verb is uniquely distinguished by the object cliticisation test, and the goal object is uniquely distinguished by obligatory marking with *?igil*.

The same possibilities of analysis are open to us for Tigre as for Dolakha Newar. Discounting the cross-linguistically implausible Analysis 1 (represented for Dolakha Newar in 28), we are left with the two possibilities we considered for Newar in the previous section, repeated here:

- (37) Analysis 2, Tigre:

|                 |                              |
|-----------------|------------------------------|
| Monotransitive: | Ditransitive:                |
| THEME           | THEME      GOAL              |
|                 |                              |
| OBJ             | OBJ      OBJ <sub>GOAL</sub> |

Analysis 2 gains most plausibility from casemarking patterns: as in Newar, obligatory marking is associated with OBJ<sub>GOAL</sub>, while marking for primary OBJ depends on topicality. Object cliticisation is harder to characterise in a non-ad-hoc manner, however: object cliticisation is possible for a theme object only in a ditransitive construction, in the presence of an OBJ<sub>GOAL</sub>.

- (38) Analysis 3, Tigre:

|                 |                          |
|-----------------|--------------------------|
| Monotransitive: | Ditransitive:            |
| THEME           | THEME      GOAL          |
|                 |                          |
| OBJ             | OBJ <sub>THEME</sub> OBJ |

Analysis 3 has the opposite problem: object cliticisation can be neatly characterised as applying to OBJ<sub>THEME</sub>, but casemarking must be characterised as obligatory for goal OBJ, and dependent on topicality for theme OBJ and OBJ<sub>THEME</sub>.

For Tigre, it may be sensible to consider a third possibility: that the two object arguments of a ‘give’-type verb both correspond to different object-like functions from the object of a monotransitive verb. This corresponds to the rare ‘tripartite’ alignment discussed by Haspelmath (2007) and Malchukov et al. (2007):

(39) Analysis 4 (tripartite), Tigre:

|                 |                      |                     |
|-----------------|----------------------|---------------------|
| Monotransitive: | Ditransitive:        |                     |
| THEME           | THEME                | GOAL                |
|                 |                      |                     |
| OBJ             | OBJ <sub>THEME</sub> | OBJ <sub>GOAL</sub> |

On this analysis, both casemarking and object cliticisation are easy to characterise: object cliticisation is possible only for OBJ<sub>THEME</sub>, and object casemarking is obligatory only for OBJ<sub>GOAL</sub>. This may provide the most satisfactory analysis of Tigre, though examination of more data may reveal additional patterns which would favor Analysis 2 or 3 over this one.

#### 9.4.3 *Applicatives: Upper Necaxa Totonac*

Beck (2006, 2007, 2008) discusses the syntactic and pragmatic properties of the applicative construction in Upper Necaxa Totonac,<sup>4</sup> showing that applicativisation promotes topical nonobject arguments to object, and that the result is a multitransitive construction in which the basic object and the applied objects share object properties.

The Totonac verb has two person agreement slots and one number agreement slot. The verb can agree with the subject only, the subject and the object, or with two objects. Subject and object agreement morphemes constitute distinct sets, with the subject affixes reflecting the person and the number of the subject, and the object affixes reflecting the person and the number of the object in the two distinct slots. The number of the object is indicated by zero in the singular and *-ka:-* in the plural, however, only one of the objects can agree for number, even if both are in the plural and agree for person.

Person agreement is obligatory with first and second person objects. If the verbs takes two first and/or second person objects, both of them control agreement. There is no subject agreement in this case.

- (40) wix kuchí:lu kín-li:-lhtuku-ya:-n      cha:-tin hótni'  
 you knife    1Obj-Instr-stab-Impf-2Obj Cl-one drunk  
 'You, knife, a drunk stabs me with you.'      (Beck 2008:2)

With third person objects, agreement is optional. Since third person singular object agreement is null, only third person plural objects are overtly marked for agreement.

<sup>4</sup>We thank David Beck for clarifying some aspects of his analysis and providing us with additional data in personal communication. Examples for which the source is not given come from personal communication with David Beck.

With monotransitive verbs, agreement with the third person plural object is very frequent, but can be absent if the object is “low” in what Beck calls “topic-worthiness”, but which we believe is better analysed as topicality. In (41), the object is nontopical, and the verb does not agree with it, showing only subject agreement:

- (41) lhú:wa' ik-lak-chukú-ma:lh                      kim-pá'hlhcha'  
 much 1SgSubj-Intensive-chop-Prog 1Poss-tomato  
 ‘I’m slicing a lot of tomatoes.’

Competition for object agreement in Totonac arises if two (or more) objects are third person: any of them can but does not have to control agreement. In (42), the verb bears two applicative affixes, and takes three object-like arguments: theme, instrumental and comitative. Beck (2008) shows that each of them can control agreement, as shown in (42a), (42b), and (42c), respectively.

- (42) a. pu:lak-kauj kin-la:xáx    na-ik-ka:-ta:-li:-tanka'  
 Class-ten 1Poss-orange Fut-1SgSubj-PIObj-Com-Instr-fell  
 kin-ta:sá:kwa wamá: hen-tin kin-machí:ta'  
 1Poss-peon this CI-one 1Poss-machete  
 ‘My peon and I will cut down ten orange trees with this machete.’
- b. hen-tu:tu'n machi':ta' na-ik-ka:-ta:-li:-tanka':  
 Class-three machete Fut-1SgSubj-PIObj-Com-Instr-fell  
 pu:lak-ti'n ki'wi' wama': chixku'  
 CI-one tree this man  
 ‘With three machetes I and this man will cut down a tree.’
- c. na-ik-ka:-ta:-li:-tanka':                      pu:lak-ti'n ki'wi' chixku'-win  
 Fut-1SgSubj-PIObj-Com-Instr-fell CI-one tree man-PI  
 kin-machi:t-kan  
 1Poss-machete-Pl.Poss  
 ‘I and the men will cut down a tree with our machete.’ (Beck 2008)

As with monotransitives, object agreement patterns depend on information structure factors. First, objects higher in animacy are more likely to control agreement than inanimate objects, although this is only a preference. More important is discourse saliency. When an inanimate object is discourse salient (i.e., topical), it becomes a legitimate controller. This is shown by the following contrast, where both (43a) and (43b) exemplify the comitative applicative.

- (43) a. \*chin-lh kin-puska':t na-ik-ka:-ta:-puzta'  
 arrive-Perf 1Poss-wife Fut-1SgSubj-PIObj-Com-look.for  
 hen-tu: kin-machi':ta'  
 Class-two 1Poss-machete  
 'My wife is here, we are going to look for my two machetes.'
- b. akxni' chin-lh kin-puska':t na-ik-ka:-ta:-puzta'  
 when arrive-Perf 1Poss-wife Fut-1SgSubj-PIObj-Com-look.for  
 hen-tu: kin-machi':ta'  
 Class-two 1Poss-machete  
 '[My machetes got lost.] When my wife comes she and I will look  
 for them.' (Beck 2008)

In (43a) number agreement with the inanimate theme object is impossible because it is not sufficiently salient. But agreement becomes acceptable if the context is modified: in (43b) the context establishes the topical role for the referent of 'machetes'. The second sentence is "about" machetes, so the object pronoun referring to machetes can control agreement. Similarly, agreement is possible with contrastive topics:

- (44) a. cha:-ti'n chixku' a'sta hen-tu: kuchi'lu li:-lhtuku'-lh ho'tni'  
 Cl-one man even Class-two knife Instr-stab-Perf drunk  
 'The drunk stabbed a man with two knives.'
- b. a'sta hen-tu: kuchi'lu cha:-ti'n chixku' ka:-li:-lhtuku'-lh  
 even Class-two knife Cl-one man PIObj-Instr-stab-Perf  
 ho'tni'  
 drunk  
 'With two knives the drunk stabbed a man.' (Beck 2008)

In (44) agreement is triggered by the instrumental topical object.

Beck (2008) proposes that the rules of object agreement obey the hierarchy in (45):

- (45) 1, 2 > discourse-salient 3 > animate 3 > inanimate 3

We believe that this hierarchy is best reformulated as the following generalisation: agreement is obligatory with first and second person objects and with topical third person objects. The correlation with animacy is only a tendency and is independently motivated: as mentioned in Chapter 3, animate referents are frequent topics. Nontopical third person objects do not trigger agreement. Note that the semantic role of the object plays no role in the object agreement pattern.

For multitransitive constructions, Beck concludes that the grammatical function of object is not unique in Upper Necaxa Totonac: he suggests that in multitransitive clauses all objects are of a similar grammatical status because any object can potentially control agreement. This implies that object agreement is independent of grammatical role.

We believe that a more satisfactory analysis is available. Bresnan and Moshi (1990) discuss object symmetries and asymmetries in Bantu, distinguishing two types of languages. In the **asymmetrical object type**, only one of a verb's complements can display primary object properties, while in the **symmetrical object type**, any of several of the verb's complements is eligible for the primary object role. They show that the distinction can be formally modelled in a very simple way in terms of intrinsic assignments of grammatical functions to arrays of semantic roles: the asymmetrical object type disallows the intrinsic association of an array of semantic roles with two unrestricted grammatical functions, while symmetric object type languages do not have this restriction. We can analyse Totonac as a symmetrical object language in their terms. Taking object agreement to be a primary object property, this means that any of several arguments can assume the primary object role and control verb agreement, though only one argument at a time can do this, since in any particular sentence there is only one primary object.<sup>5</sup>

This view comports well with our view that information structure can play a crucial role in the mapping of semantic roles to grammatical functions. We suggest that secondary agreement in Upper Necaxa Totonac is triggered by the primary OBJ, which can correspond to a variety of semantic roles. In contrast, secondary objects OBJ<sub>θ</sub> do not trigger agreement. In monotransitive constructions, the choice between OBJ and OBJ<sub>θ</sub> is determined by the referential status of the object (first or second person vs. third) and, for the third person object, topicality. Applicative operations produce clauses with up to five objects, but if they are all third person, only one argument controls verb agreement: the agreeing object is the primary OBJ. Non-agreeing objects are semantically restricted and correspond to OBJ<sub>THEME</sub>, OBJ<sub>GOAL</sub>, OBJ<sub>BENEFACTIVE</sub>, OBJ<sub>INSTRUMENTAL</sub> and so on. The semantic type of each secondary object is indicated by applicative morphology on the verb, except for OBJ<sub>THEME</sub>, which does not require morphological applicativisation. This analysis makes Upper Necaxa Totonac essentially similar to Ostyak: in both languages, agreement is defined by grammatical functions. Both languages impose a strict alignment between grammatical functions and topicality: the primary object is topical, while the secondary object is a nontopic. The difference between the two languages lies in the number of semantically restricted secondary objects in the

<sup>5</sup>MacKay and Trechsel (2008) argue that the closely related language Misantla Totonac is a symmetrical language in the sense of Bresnan and Moshi (1990).

inventory of grammatical functions: Ostyak only has  $OBJ_{THEME}$  and expresses other semantic roles by means of casemarked obliques, whereas Upper Necaxa Totonac allows a number of semantically restricted objects with various functions and therefore allows for a variety of multi-object constructions. However, since there is tight alignment between marking and grammatical function, “doubling” of grammatical marking of the kind observed in Tigre, Newar and non-standard Hindi is impossible for Upper Necaxa Totonac.

Beck (2007) considers a potential object test that distinguishes the theme/basic object of the verb from applied objects: the theme is the only object that can be targeted by the “object suppression” voice, which blocks the syntactic expression of the object. Object suppression cannot target the applied object. According to Beck (2007), object suppression is not a good test of primary objecthood, since verbs in the object suppression voice can still agree with applied objects. Under the assumption that object suppression makes the basic/theme object syntactically inactive, and given Bresnan and Moshi’s theory of symmetrical vs. asymmetrical object languages, it is not surprising that verbs in the object suppression voice can show object agreement: when the theme object is suppressed, the applied object can fill the primary  $OBJ$  role. However, the existence of this test leads to a consideration of an alternative analysis of the data: it is possible that object suppression does indeed target the primary, non-applied object, and that verb agreement in Totonac can be controlled by any topical object argument, whether  $OBJ$  or  $OBJ_{\theta}$ . We have not encountered other languages exhibiting this pattern, and it goes against our claim that topical marking is associated with grammatical functions that are high on the grammatical hierarchy (for objects,  $OBJ$  vs.  $OBJ_{\theta}$ ). However, the possibility cannot be ruled out in principle; further exploration of object patterns in Totonac may provide evidence indicating whether this alternative analysis is viable.

### 9.5 Topicality and goals

We have seen in this chapter that topical patient/themes often share properties in common with goals of ‘give’-type verbs; these properties may be limited to marking, or may extend to grammatical role. The patient/theme object marker is often identical to the marker of the goal argument in constructions with ‘give’-type verbs, independent of grammatical function. For instance, in non-standard Hindi and possibly in Dolakha Newar and Tigre, the theme and goal correspond to different grammatical functions, but still bear the same grammatical marking. In languages where the secundative/double object construction is possible, the similarities go further: in standard Hindi, Chatino and some constructions in Ostyak the goal is not only encoded identically to the topical theme object, but represents the same grammatical function.



This is a cross-linguistically widespread phenomenon and has often been noted in previous work. Bossong (1991:157) observes that the dative marker is a common (but not the only) source of innovated accusatives in languages with DOM. Masica (1982) provides an overview of a number of languages (predominantly in Asia) where object casemarking coincides with the marking of the goal argument in ditransitive constructions. Apart from the languages discussed in this book, similar patterns are observed in Romance and Semitic. In Klamath (Penutian, Rude 1982) patient/theme objects are either marked or unmarked for case, while for goal and beneficiary objects the same marking is obligatory. Similar polyfunctionality is present in Lezgian (Nakh-Daghestanian) and Imonda (Papuan) (Heine and Kuteva 2002:38) and many languages cited by Kittilä (2006a,b). Languages with object agreement often show the same pattern, e.g. Palauan (Austronesian), Huichol (Uto-Aztecan) and a number of other languages discussed by Georgopoulos (1992).

Kittilä (2006a) suggests that these marking patterns follow from the fact that, on the one hand, DOM depends on animacy and, on the other hand, there is a strong correlation between animacy and recipient (goal) (see also Haspelmath 2007). Recall that for Kittilä, “animacy” is a cover term for topic-worthiness or topicality, including definiteness and prominence as well as animacy. In fact, ditransitive goals are overwhelmingly high in topic-worthy features, but in our view animacy effects are subsidiary and follow from independent requirements on information structuring. In particular, we follow previous research in analysing goals in ‘give’-type constructions as unmarked secondary topics.

Dryer (1986), Polinsky (1998) and Haspelmath (2004) claim that ditransitive goals are inherently more topical than theme arguments. According to Polinsky (1998), the goal is associated with a pragmatic presupposition of independent existence at least prior to the event of transfer, and does not require assertion in the sentence where it appears. In contrast, the existence of the patient is not necessarily presupposed; this matches the prototypical properties of the theme of monotransitive verbs, whose existence is not necessarily presupposed prior to the event in question.

Like primary topic, secondary topic is under discussion at the time of the utterance and is expected to carry a pragmatic presupposition of existence. This semantic difference corresponds to an asymmetry in the relative informational status of the two arguments: the goal is superior to the patient in topicality, and other things being equal, the patient is more likely to be interpreted as focus and the goal as secondary topic. In other words, secondary topic is the default information structure role for the goal in ‘give’-constructions. The same is true of the causee argument, which prototypically carries a presupposition of existence because it is indispensable for the development of the causative event (Polinsky 1995).

Since the ditransitive goal is a frequent secondary topic, it is likely to bear secondary topic marking. We will see in Chapter 10 that in Iranian and Indo-Aryan DOM started as secondary topic marking on goals but later spread to topical patient/theme objects, due to the fact that objects are more often topical than other nonsubject grammatical functions. Thus, the uniformity of grammatical marking on topical monotransitive objects and ditransitive goals observed in some languages has a historical explanation, even though, as we have seen in this chapter, the grammatical functions associated with these arguments are not necessarily the same.

## 9.6 Conclusion

We have seen that languages can express the goal and theme arguments of 'give'-type verbs in various ways, depending on the grammatical functions which the language can deploy, the mapping rules relating grammatical functions to semantic roles, constraints on marking possibilities, and the availability of double object constructions. These are, in principle, independent factors.

In languages of the type discussed in Chapter 8, grammatically marked objects correspond to OBJ and are typically topical, while grammatically unmarked objects correspond to OBJ<sub>θ</sub> and are typically nontopical. In monotransitive constructions, the complement is realised as OBJ if topical, and OBJ<sub>θ</sub> if nontopical. In languages without a ditransitive construction where the goal argument of a 'give'-type verb is expressed as an oblique, possibilities for expression of the theme argument are exactly as for monotransitives, since the presence of an oblique goal does not constrain the marking or grammatical function of the theme object.

In multitransitive constructions, when more than one nonsubject argument is present, mapping possibilities may be restricted by the inventory of restricted/secondary objects in the language, in which case the correlation between marking and information structure is obscured. In Hindi and Chatino, the goal argument of a ditransitive verb is the primary object (OBJ) and the theme is the secondary object (OBJ<sub>THEME</sub>); the theme must be unmarked and the goal must be marked, independently of their information structure role. This means that there is no DOM in ditransitive constructions: the theme and the goal correspond to two different object functions and are associated with different grammatical marking. On the other hand, in Tigre and Dholaka Newar, identical marking of the two objects is possible because grammatical marking and grammatical functions are not in tight correspondence.

Data from all of the languages we have examined in this chapter reinforce our basic claim: objecthood, and particularly primary objecthood, is inherently associated with topicality.

## *Semantic features, topicality and grammaticalisation*

As discussed in Chapter 1, most previous work on DOM is based on semantic features such as person, animacy, definiteness and specificity (Diesing 1992, van Geenhoven 1998, Ritter and Rosen 2001). However, we have seen that these features are not enough to explain object marking patterns in many languages, and we have argued that in many cases the appropriate generalisations involve the notion of topicality, especially secondary topicality. The topicality of a referent depends on how the speaker construes the situation within the given communicative context; features of topic-worthiness determine only the **likelihood** for the object to be construed as topical.

What our account still must explain is why in some languages DOM does in fact depend on semantic features of topic-worthiness rather than the information-structural notion of topicality, sometimes in combination with information structure and sometimes independent of it. In this chapter we propose a historical explanation for these patterns, based on the notion of grammaticalisation. We suggest that semantic patterns of DOM can arise as a result of different stages and directions of grammaticalisation of topical marking on objects.

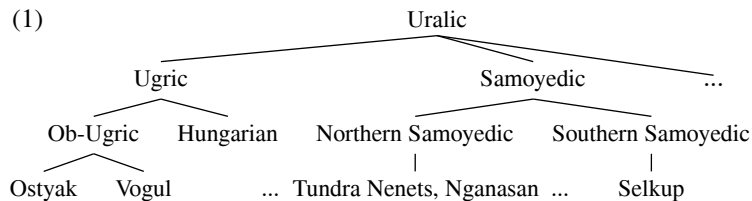
### **10.1 Case studies**

We sketch the historical evidence for the emergence of DOM in Uralic, Persian and Hindi. Presenting a fuller diachronic picture will of course require a much more detailed investigation; we view this chapter as the first step in this direction.

10.1.1 *Uralic*

In this section we examine grammaticalisation of topical marking in three genetically related Uralic languages, building on our discussion in Chapters 7 and 8 of grammatical marking via object agreement in Ostyak and Tundra Nenets, and we outline a proposal for the historical development of object marking in Uralic.

Ostyak is a member of the Ugric branch of Uralic. Tundra Nenets is a member of the Samoyedic branch of Uralic, and is therefore distantly related to Ostyak. Besides Ostyak (manifested in a number of dialects), the Ugric branch includes Ostyak's closest linguistic relative Vogul (or Mansi); these form the Ob-Ugric subgroup of Ugric, which also includes the more distantly related language Hungarian. The Samoyedic group consists of the Northern Samoyedic subgroup (Tundra Nenets, Forest Nenets, Tundra Enets, Forest Enets and Nganasan) and the Southern Samoyedic subgroup, whose only living representative is Selkup.



Importantly, the Samoyedic and Ugric languages are geographically close and demonstrate a number of common features (Xelimskij 1982). They are sometimes jointly referred to as Eastern Uralic languages.

Xelimskij (1982) and others have shown that object agreement in Ugric and Samoyedic has a common origin (although it was probably absent in Western Uralic): at least some agreement markers go back to the same etymological source. It is also likely that conditions on agreement were identical at some stage. However, conditions on agreement differ in the modern Ugric and Samoyedic languages.

We suggest that the Ostyak system of DOM based solely on information structure is the most archaic and can probably be hypothesised for Proto-Eastern-Uralic — that is, those Proto-Uralic dialects from which the Ugric and Samoyedic languages developed. We discuss the situation in Samoyedic below. For the Ob-Ugric branch, this claim is supported by the fact that apart from Ostyak, the information structure-driven system is found in the related language Vogul (Skribnik 2001). In this language, object agreement works similarly to Ostyak. It does not directly depend on definiteness, as shown by the examples in (2): in (2a) the object is definite, but there is no agreement, and in (2b) the object is indefinite but agrees.

- (2) a.  $\chi\bar{u}rum\ l\bar{u}nt\ jox\bar{t}\text{-}s\text{-}\bar{e}t$                        $os\ tawe\ kins\text{-}\bar{e}\gamma\text{-}\bar{e}t$   
 three goose come-Past-3PISubj and he.Acc search-Pres-3PISubj  
 ‘Three wild geese appeared and are looking for him.’
- b.  $j\bar{u}swoj\text{-}t\ t\bar{a}n\ os\ m\bar{a}n\ p\bar{a}si\gamma\text{-}kwe\text{-}t$                        $now\text{-}i\text{-}jan\bar{e}l$   
 eagle-Pl they and little reindeer-Dim-Pl catch-Pres-PlObj.3PISubj  
 ‘Eagles, they hunt little reindeer, too.’                      (Skribnik 2001:226)

Skribnik (2001) argues that agreement is governed by the topicality of the direct object. As in Ostyak, agreeing objects in Vogul can have a variety of semantic roles. For instance, in (3) the object corresponds to a goal or beneficiary (the doll), which is clearly topical given the preceding context and corresponds to a referential null.

- (3) [The Mos-woman has something that looks like a small child: a doll.]

$sa\gamma i\text{-}l\ w\bar{a}r\text{-}i\text{-}te,$                                                $w\bar{a}i\text{-}l$   
 coat-Instr make-Pres-SgObj.3SgSubj shoe-Instr  
 $w\bar{a}r\text{-}i\text{-}te$   
 make-Pres-SgObj.3SgSubj

‘She is making a coat for it, she is making shoes for it.’

(Skribnik 2001:229)

In fact, topicalisation by means of object agreement goes even further in this language, because a wider variety of semantic roles than in Ostyak can be realised as an agreeing object. Topicalised and nontopicalised location arguments are presented in (4), whereas in (5) we show topicalised and nontopicalised instruments.

- (4) a.  $taw\ \chi\bar{a}p\text{-}\bar{e}n\ jox\bar{t}\text{-}\bar{e}s$   
 he boat-Dat come-Past.Subj.3Sg  
 ‘He came to the boat.’
- b.  $taw\ \chi\bar{a}p\ jox\bar{t}\text{-}\bar{e}s\text{-}te$   
 he boat come-Past-Obj.Sg.Subj.3Sg  
 ‘He reached the boat.’                                              (Skribnik 2001:229)
- (5) a.  $am\ tul'\bar{o}wl\text{-}um\text{-}\bar{e}l\ r\bar{a}ta\bar{s}l\text{-}\bar{e}\gamma\text{-}um$   
 I finger-1Sg-Instr tap-Pres-1SgSubj  
 ‘I am tapping with my finger.’

- b. am tul'owl-um rātašl-i-l-um  
 I finger-1Sg tap-Pres-SgObj-1SgSubj  
 'I am using my finger for tapping.' (Skribnik 2001:229)

In both cases a topical element may not appear as an oblique element, but must be promoted to object (on our analysis, primary object) and must agree with the verb. In Ostyak, the primary object can correspond to a patient/theme, goal, causee or beneficiary, while in Vogul objects can also bear other semantic roles such as location and instrument.

Thus, in both Ob-Ugric languages, Ostyak and Vogul, object agreement is conditioned by information structure, and we hypothesise that this was also true in Proto-Ob-Ugric. There are no known semantic restrictions on agreeing objects in Ob-Ugric.<sup>1</sup>

The situation in the third Ugric language, Hungarian, is different. First and second person object pronouns in Hungarian never trigger agreement; this is similar to the Nenets pattern discussed in Chapter 7, Section 7.2:

- (6) lát-nak / \*lát-ják téged / engem  
 see-3PISubj see-Obj.3PISubj you me  
 'They see you/me.' (É. Kiss 2002:54)

Unlike Nenets, however, third person agreement is not determined by information structure. Instead, the triggering feature is definiteness. Nouns preceded by the definite article *a/az*, possessed nouns, proper nouns, complement clauses, third person personal pronouns, and a number of other types of pronouns, including reflexive, reciprocal, and some interrogative pronouns, all count as definite objects and trigger object agreement. Indefinite nouns preceded by the indefinite article *egy* or quantifiers, as well as bare objects, do not trigger agreement.

- (7) a. János zöldre fest-ett-e / \*fest-ett a  
 J. green paint-Past-Obj.3SgSubj paint-Past.3SgSubj Det  
 kapu-t  
 gate-Acc  
 'János painted the gate green.' (É. Kiss 2002:70)
- b. Bicikli-t sok lány lát-ott / \*lát-ta  
 bicycle-Acc many girl see-Past.3SgSubj see-Past.Obj.3SgSubj  
 'Bicycle, many girls saw.' (É. Kiss 2002:22)

<sup>1</sup>However, in some Eastern varieties of Ostyak the distribution of object agreement is similar to Nenets: agreement is impossible for first and second person objects.



was the triggering feature for agreement in early Hungarian: although the objects in (10a,b) are definite, they are unlikely to be topical because these objects introduced novel discourse participants.

- (10) a. an-nak szabadás-á-t én szív-em kíván  
 that-Dat liberation-3Sg-Acc I heart-1Sg desire.3SgSubj  
 ‘My heart desires his liberation.’ (Marcantonio 1985:290)
- b. es tarsibeli kiralog es zigetek aiandok-a-i-t aiandokoz-nak  
 and friendly kings and islanders present-3-Pl-Acc give-3PlSubj  
 ‘And friendly kings and islanders give their presents.’  
 (Marcantonio 1985:290)

Object complement clauses did not necessarily trigger agreement either. Marcantonio (1985) argues that such examples can be found even in more recent Hungarian literature, especially in poetry. They show that the rules of object agreement in earlier Hungarian deviate from modern Hungarian and were likely to be more dependent on information structure.

An information-structure-based system of object agreement is also present in Samoyedic, as shown by the Tundra Nenets data presented in Chapter 7, Section 7.2, although Tundra Nenets stands somewhat “in-between” Hungarian and Ob-Ugric: object agreement marks topical arguments, as in Ostyak, but first and second person pronouns never agree, as in Hungarian. The other Samoyedic languages, including Selkup and Nganasan, are similar to Nenets in this respect. In these languages object agreement is only possible with 3rd person objects, as shown by the Selkup examples in (11):

- (11) a. Təp šjnty qontyrəntə / \*qontyrəntyjty  
 he you.Acc see.Fut.3SgSubj see.Fut.Obj.3SgSubj  
 ‘He will see you.’ (Kuznecova et al. 1982:235)
- b. Təp kanap qontyrəntə / qontyrəntyjty  
 he dog.Acc see.Fut.3SgSubj see.Fut.Obj.3SgSubj  
 ‘He will see a/the dog.’ (Kuznecova et al. 1982:235)

According to Kuznecova et al. (1982), object agreement marks definite third person objects, although the examples they provide do not demonstrate clearly whether we are dealing with definiteness or topicality. If it is indeed definiteness, Selkup is just like modern Hungarian, except that it does not have grammaticalised expression of definiteness in the form of articles. For Nganasan, Tereshchenko (1979) explicitly says that object agreement with third person



objects depends on “logical emphasis”, which is her terminology for information structuring, whereas definiteness does not play a decisive role. For instance, in both examples in (12) the object is definite (possessed); Kuznecova et al. (1982) state that the object in (12a) is more salient than in (12b), and agreement is required only in (12a).

- (12) a. Bukurjakumti mýtəmi”əgəj byzədja najbəgə  
 net.Acc.Sg.3Du pushed.Obj.3DuSubj water.Dat long  
 nirkutənu  
 willow.branch.Instr  
 ‘They (dual) pushed their net into water with long willow branches.’  
 (Tereshchenko 1979:190)
- b. Bukurjakumti najbəgə nirkutənu byzədja  
 net.Acc.Sg.3Du long willow.branch.Instr water.Dat  
 mýtəmi”əzi  
 pushed.3DuSubj  
 ‘They (dual) pushed their net into water with long willow branches.’  
 (Tereshchenko 1979:191)

In Samoyedic and Old Hungarian, then, object agreement was restricted to third person topical objects. In other words, the scope of grammatical marking of topics was reduced. This situation contrasts with Ob-Ugric, where agreement marks topical objects independently of person.

Comrie (1977:10) suggests a functional explanation for the lack of agreement with first and second person controllers: first and second person pronouns are inherently definite, so there is no need to mark them explicitly. This explanation is based on the premise that the primary function of object agreement in Eastern Uralic is the marking of definiteness. However, in all Eastern Uralic languages except modern Hungarian, information structure and not definiteness plays the primary role in patterns of object agreement. The Hungarian situation thus is likely to be secondary, as is also confirmed by the Old Hungarian data.

An alternative explanation for first/second person agreement in Uralic is as follows. The referents of the first and second person pronouns are highly salient in human communication. In fact, they occupy the highest position on scales of topic-worthiness, and are the most likely topics. However, the primary topic is more salient than the secondary topic by definition, so first and second person pronouns are more likely to correspond to the primary topic than the secondary topic. Given the default alignment between the primary topic and the subject, first and second person pronouns tend to be encoded

as subjects: in fact, in Ostyak they rarely occur as topical objects triggering agreement. If they correspond to the patient/theme of a transitive verb, passivisation is preferred even in contexts which in principle require secondary topicalisation of the object. For example, in the context ‘Where did John hit Peter?’ the answer (given in 13) will invariably be in the active, and the secondary topic object ‘Peter’ will trigger object agreement. But in the context ‘Where did John hit you?’ there is room for variation: some speakers prefer to use the active construction with object agreement (example 14a), while other speakers employ passivisation (example 14b).

(13) Juwan Pe:tra xo:t-na re:skə-s-li  
 John Peter house-Loc hit-Past-Obj.3SgSubj  
 ‘John hit Peter in the house.’

(14) a. Juwan mane:m xo:t-na re:skə-s-li  
 John I.Acc house-Loc hit-Past-Obj.3SgSubj  
 ‘John hit me in the house.’

b. (ma) Juwan-na xo:t-na re:skə-s-aj-əm  
 I John-Loc house-Loc hit-Past-Pas-1SgSubj  
 ‘I was hit by John in the house.’

Presumably the choice depends on the individual assessment of the saliency of relevant referents, but the point is that there is an asymmetry between (13) and (14).

On this view, the Samoyedic languages (Nenets, Selkup and Nganasan) and Old Hungarian have grammaticalised the tendency for first and second person pronouns to be likely primary topics and unlikely secondary topics.<sup>2</sup> Therefore they cannot correspond to the primary object, which is strongly aligned with secondary topic in these languages. There are no such restrictions for third person objects. The next historical stage is represented by Hungarian (and possibly Selkup): grammatical marking of third person topical objects extends to nontopical definite objects. This process illustrates the spreading of grammatical marking to nontopical objects bearing topic-worthy features, and the concomitant loss of a connection to information structure.

<sup>2</sup>This tendency has been discussed elsewhere in the literature. For instance, Schulz (2005) argues, based on facts from German pro-drop, that unmarked topical objects are third person, not first or second person. There is no such asymmetry for subject topics. Haspelmath (2004) cites frequency data indicating that first and second person themes are less common and less prototypical than third person themes, both in monotransitive and ditransitive constructions, and that in some languages they cannot map as objects at all.

## 10.1.2 Persian and the Iranian languages

DOM exists to some degree in almost all Iranian languages except for Pashto, where the object is never marked, and Kurdish, where it is always marked by the accusative case. Conditions on DOM differ across languages. Moreover, according to Bossong (1985), there are several etymological sources of DOM in Iranian. In this section we summarise the history of Persian *râ* based on the work of Bossong (1985:58ff.), Karimi (1990), Dabir-Moghaddam (1992), Haig (2008), and Paul (2003).

Persian *râ* has cognates in Balochi, Gazi, Tat, Luri and a number of other Iranian languages. Phonologically, this item has undergone a reduction from the Old Persian *râdiy* (6th century BC to 3rd century BC) to Middle Persian *rây* (ca. 225 AD to 651 AD) and New Persian *râ*. Semantically and functionally, it has undergone a change of status from a noun to a postposition with mixed semantic and syntactic functions, then to a (secondary) topic marker restricted to certain grammatical functions, and then established as an object marker several centuries after the loss of the old Proto-Iranian oblique case used for direct objects.

The ancestor of *râ* was a noun meaning something like ‘reason, aim’. It was attested with the postpositional meaning ‘for the sake of, on account of, by, due to, because of’ in Old Persian. Later, in Middle Persian, this postposition came to be used as a marker of the beneficiary, purpose, goal, external possessor and a few other oblique functions. We can see the reflexes of this situation in modern Persian as well. As was shown in Chapter 6, Section 6.1.1, *râ* can mark the external possessor in modern Persian. However, not all of its earlier functions are preserved. For instance, Bossong (1985) cites the following Middle Persian sentence:

- (15) ma-ra dar šahr dust-ân besyâr-and  
 I-RA in town friend-Pl many-3Pl  
 ‘I have many friends in town.’ (Bossong 1985:61)

In this example *râ* marks the predicative external possessor, but modern Persian employs the verb ‘have’ in predicative possessive constructions.

In the Early New Persian period, also called Classical Persian (around 1000 AD), *râ* was mainly used on indirect objects. Its use on direct objects appears to be an innovation, though sporadic cases of *râ* on definite direct objects were attested in the 10th century. In the Classical Persian example below *râ* is rendered orthographically as *r’*.

- (16) k=š ’yn d’t’n r’ bgft  
 when-3Sg this story RA tell.Past.Irr  
 ‘when he told this story’ (Haig 2008:128)

However, most objects remained unmarked even if they were highly definite:

- (17) *xosro tus be du dâd*  
 king Tus to him gave  
 ‘The king gave Tus to him.’ (Karimi 1990:110)

The equivalent of (17) would be ungrammatical in standard Modern Persian, although it may be acceptable in contemporary dialects. In late New Persian, *râ* completely lost its indirect object function and came to be an obligatory marker for definite direct objects and some non-object topical elements (see Chapter 6, Section 6.1.1), though its earlier oblique function is still preserved in some lexical items such as ‘why?’.

We hypothesise that *râ* on objects originates as a topicality marking device. As discussed in Chapter 9, beneficiary and goal arguments are unmarked secondary topics, and external possessors are also highly topical (Payne and Barshi 1999). Before it began appearing on objects, *râ* was systematically used in these two functions. Because of the close association of these functions with topicality, *râ* is likely to have been reanalysed as a topicality marker instead of a marker of a grammatical function. It subsequently began appearing on topical objects, giving rise to DOM.

Interestingly, *râ* was first attested on objects high in topic-worthiness. Haig (2008:152-153) argues that as an object marker *râ* first appeared on first and second person pronouns (as is still observed in some related Iranian languages), and was then extended to lexical nouns. According to Paul (2003:182), in the Early New Persian period it was mostly found on animate definite nouns, while inanimate definites remained unmarked. So although the syntactic functions of *râ* were reduced from the Old Persian period, its use as an object marker can be characterised as a continuous expansion: it first appeared on first and second person pronouns, then spread to animate definite objects, then all definite objects, but was still optional. Later, in modern Persian, it became obligatory on all definite objects and topical indefinite objects.

### 10.1.3 *Hindi and the Indo-Aryan languages*

The Indo-Aryan languages are distantly related to the Iranian languages, together constituting the Indo-Iranian branch of Indo-European. The Indo-Aryan languages, including Hindi, are descended from Sanskrit; the Old Indo-Aryan period, encompassing Vedic and Classical Sanskrit, extended from 1500 BC to 600 BC, and was followed by Middle Indo-Aryan, which extended from 600 BC to 1000 AD (Masica 1991). Middle Indo-Aryan lost almost all case inflections which had been present in Sanskrit, including the original accusative

which merged with the nominative, but the new Indo-Aryan languages developed new postpositional or clitic-like markers of major grammatical functions. These typically originated from lexical nouns and verbs.

According to Beames (1966 [1872-79]), Hindi *ko* goes back to the Sanskrit noun *kaksha* ‘armpit, side’. As argued by Butt (2008a) and Butt et al. (2008), the original function of this element was purely locational: the word meaning ‘armpit’ grammaticalised as a spatial postposition. Indeed, in related Iranian languages its cognate denotes location. For instance, the Iranian language Pashto has a locative in *kii/ke*, which goes back to Avestan *kaaše*. The latter is the locative form of *kaaša* ‘armpit’, etymologically related to Sanskrit *kaksha* (Hewson and Bubenik 2006:150).

The early Hindi forms of this postposition, *kaham*, *kum*, *kaum*, *kau* etc., all reflect the Old Hindi accusative form *kākha*. Following further phonological changes, it developed into Hindi *ko*, Bengali *ke*, Oriya *ku*, Singhi *khe* and Siraiiki *kon* (Masica 1982). In spite of their common origin, the exact conditions on the object-marking use of this element differ in modern Indo-Aryan languages. For instance, the range of functions of the object marker in Bengali is narrower than in Hindi: Bengali *ke* is mostly used on animate objects, and is not compatible with pronouns referring to things.

The oldest documented examples of *ko* in Hindi come from the early 13th century (Beames 1966 [1872-79], Kellogg 1955 [1893]). Beames examines the early occurrences of this postposition, which at this stage signalled the purposive and the recipient or goal of ditransitive verbs like ‘give’. There are also a few examples where *ko* marks objects. In most cases, such objects can be interpreted as unattained goals/endpoints, for instance with the verb ‘seek’. Such examples are not numerous, however. According to Butt (2008a), this usage continues over centuries, with a handful of locative/possessive uses documented in the 1500s. The following example represents Divan-e-Hasan Shauqi (Deccani Urdu of 1564).

- (18) na band<sup>h</sup>iya kad<sup>h</sup>i zara ōni peṭ kō  
not tie.Perf.MascSg ever armour Pron.3.Gen belly Obj  
‘He never tied armour to/on his belly.’ (Butt 2008a:16)

In the 1800s, *ko* appears with verbs of directed motion such as ‘go’ or ‘reach’.

- (19) is manzil-ko kab poāco-ge  
that destination-Obj when reach-2Sg-Fut  
‘When will you reach that destination?’ (Butt 2008a:17)

It is also well known that *ko* marks experiencer subjects and the subjects of modal verbs expressing obligation (Mohanani 1994, Butt 2008a), although this usage seems to be fairly recent.

Butt (2008a) presents the following hypothesis for the development of the functions of *ko* in Hindi/Urdu. After this element was drawn into the system of spatial postpositions, its development went in two directions. On one hand, it developed a dynamic meaning denoting the endpoint of a situation ('to') and, on the other hand, it acquired a stative meaning ('at'). The next step was a metaphorical extension by which the spatial concepts were reanalysed in the domain of events and participants. The dynamic final point interpretation of *ko* was extended to mark participants as being associated with the final part of an event, in particular, recipient/beneficiary arguments of ditransitive verbs. Such arguments are often thought of as abstract goals. Animate locations ('at') were reanalysed as subjects (Butt et al. 2006).

The next question is how the spatial/goal *ko* became an object marker. This development is not immediately obvious (cf. Ahmed 2006). A number of authors relate the accusative usage of *ko* to its spatial meaning through a type of metaphoric extension. For instance, Mohanan (1994) argues that accusative marking in Hindi is used for logical objects towards which an action or event is directed. That is, it can be seen to mark the endpoint or the goal of a (bound) action. According to Butt (2008a), in the modern language (the Urdu variety of Hindi/Urdu) *ko* marks specific objects. This meaning derives from its function to express (not necessarily attained) endpoints that are abstract but specific. Roughly speaking, then, the objective marking has its roots in spatial marking, and objects are reanalysed endpoints.

We do not disagree with the essence of this explanation, but believe that additional factors may play a role.

Importantly, Hindi is not alone. Heine and Kuteva (2002) show in their substantial survey of grammaticalisation paths that allative and dative are the two most common sources of object marking, and we demonstrated in previous chapters that in many other languages the goal argument of a verb like 'give' is grammatically marked in the same way as the transitive patient/theme. However, not all these markers have a spatial origin. As we saw in the previous subsection, the Persian object marker developed from the marker for the ditransitive goal, but the latter does not go back to a locative noun (its source is a noun meaning something like 'reason'). Deo (2008) analyses the origin of the accusative-dative *-lā* in Modern Marathi, tracing it back to the Old Marathi adposition *lāgim*, which expressed purpose or benefaction. This postposition is attested in Middle Marathi poetry, but always with a purely benefactive/purposive use. It did not indicate spatial proximity or control over the situation. In the late nineteenth and early twentieth century it emerged as the default dative case marker in the language, and in modern Marathi it is used as the differential object marker. Interestingly, this object marker replaced the already existing accusative marker *-si*: that is, the dative marker *-lā* replaced the existing device for accusative marking rather than filling a gap. Thus, evidence

from Marathi seems to suggest that dative marking tends to be diachronically extended to accusative contexts even when it does not have a spatial origin, and despite the presence of a morphologically distinct accusative. Even more striking is that accusative-dative homophony exists in languages where it is marked by agreement, which can hardly be thought of as having a spatial origin. Rather, we claim that it originates from the need to keep track of referents and often cross-references topics (see Section 10.2.1 of this chapter). In sum, we believe that there must be more to the accusative-dative connection than the idea that objects are endpoints reanalysed from the spatial domain to a more abstract domain. Also recall that, on our account, specificity as such does not explain the distribution of objective *ko* in modern Hindi (the situation may be different in modern Urdu). As we argued in Chapter 8, nonspecific objects are unmarked, but specific inanimates are either marked or not, depending on topicality. This implies that the explanation for the accusative-dative homophony may have to do with information structuring.

We believe that when *ko* started to function as the regular marker of direct objects which could co-occur with any transitive verb, it first marked topical objects. According to Masica (1982), *ko* became established as a general object marker fairly late, appearing only once in the early prose of Braj Bhasha, the literary form of Medieval Hindi (17th century). At that stage it was mostly found on pronominal objects, although it is difficult to judge how grammaticalised it was in this function. Non-human objects were very rarely marked, and even definite human patient/theme objects could remain unmarked, as shown by the following examples from medieval Hindi:

- (20) a. haraṣe Lakhana dekhi dou bhrātā  
rejoice Lakshman see two brothers  
'Lakshman rejoiced to see his two brothers.' (Masica 1982:44)
- b. ma-tā Bharatu goda baiṭha-re  
mother Bharatu lap seated  
'The mother seated Bharatu in her lap.' (Masica 1982:19)

Specific human objects are not obligatorily marked with *ko* even in later Hindi, including the twentieth-century examples in (21) and (22):

- (21) a. [Context: I am a sinner before God, Professor Saheb!]
- b. māī-ne baccaa badal diyaa  
I-Erg children switch give.Past  
'I switched children.' (Masica 1982:20)

- (22) *brīṭen ke mantrimañḍal ne apne tiin pratinidhi bhaarat bheje*  
 Britain Poss cabinet Erg its three representatives India send  
 the  
 Past

‘Britain’s Cabinet had sent its three representatives to India.’

(Masica 1982:19)

Example (21) is from a novel by Rangeya Radhava, written in 1961, and example (22) is from a text by Yash Pal, written in 1963. The objects in (21) and (22) are likely to inhabit the focus domain and be nontopical. Similar examples are cited by Kellogg (1955 [1893]) and McGregor (1972).

In sum, evidence indicates that earlier varieties of Hindi may have patterned differently from the modern language as far as object marking is concerned. The distribution of *ko* as an object marker was much more restricted: it was not required on all specific human objects, as it is in the modern language (McGregor 1972:185). We suspect that, as in Persian, the Hindi object marker originated as a marker of secondary topicality on ditransitive goals (prototypical secondary topics). Later, it started to be used to indicate topical patient/theme arguments, being mostly used for highly likely topics (personal pronouns). From this function it expanded to mark specific human objects independently of their information structure function, reflecting the fact that topics tend to be human/animate. As in Persian, then, the role of the referential semantic features of the object has increased, and the connection to information structure has weakened.

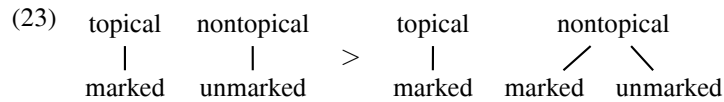
## 10.2 Paths of grammaticalisation

Based on the evidence discussed in the previous section, we argue that DOM first emerges as an information structure marking device, at least in some languages. Only later does it come to depend partially or completely on referential features of the object. In its essence this idea is similar to Danon’s (2006:1005) suggestion that “DOM might initially arise out of functional factors, and later, as grammaticalisation proceeds, become syntactically governed”. The difference is that Danon includes referential semantic features together with pragmatic and cognitive considerations in the list of possible functional factors, while on our account functional factors are limited to information structure alone.

### 10.2.1 *Spreading of DOM*

One possible direction of change involves widening of grammatical marking, where topical marking of objects extends to certain nontopical objects. This process can be schematically represented as follows:





That predicate-argument agreement often originates as a topic-marking device is well known. The diachronic connection between free-standing pronominal topics, incorporated pronouns or clitics, and agreement affixes has been much discussed since at least Givón (1979). For instance, Bresnan and Mchombo (1987:777) argue that object marking in Bantu emerged as anaphoric topic marking and that this function is still active in Chicheŵa, as we saw in Chapter 2. In Chicheŵa, object markers are incorporated pronouns referring to topics. In other Bantu languages, object markers have partially or totally lost their pronominal reference and have undergone grammaticalisation as agreement morphemes; in some Bantu languages, agreement applies unexceptionally either to all objects or to certain semantic classes of objects, independently of their topicality.

Morimoto (2002) presents a historical view of object agreement marking in Bantu that is in many respects compatible with our view: object agreement originates as a topic-marking device, and later comes to mark certain semantic features associated with topicality, such as definiteness. When incorporated pronouns develop into full-fledged grammatical agreement morphemes, they become obligatory for all objects or for a subset of objects bearing relevant semantic features. Topic-anaphoricity may still be visible but, given that linguistic change is gradual, it is observed to different degrees in different Bantu languages. In addition, there is a great degree of inter-speaker variation in most languages. Such variation is expected, since synchronic variability is a sign of an unstable situation in a diachronic process of transition from pronominal incorporation to grammatical agreement. Morimoto’s account differs from ours in assuming that agreement is “optional” in some cases, in line with much OT-based work, since objects with the same semantic features either do or do not trigger agreement; she treats this “optionality” in terms of constraint reranking in an OT setting, along the lines discussed in Chapter 1, Section 1.3. On this view, the optionality of DOM is explained by the fact that transitional stages in the diachronic process exhibit different properties both across the family and within individual languages. In our account, real optionality never arises: conditions on agreement are not formulated in terms of referential semantic features of the object. Rather, they are defined in terms of the information structure role of the object (sometimes in combination with semantics) or its grammatical function, but in either case there is no variation. For instance, Ostyak objects with the same semantic features may but need not agree. However, agreement is not optional in any syntactic sense: it is triggered by primary

objects aligned with the information structure role of topic, while secondary objects aligned with nontopics do not show agreement.

Nevertheless, we agree with Morimoto that patterns of DOM in different languages reflect different stages of grammaticalisation, and that agreement arises as an indicator of topicality, and only later comes to depend on referential semantics. Her analysis and ours agree on the direction of the grammaticalisation process: topical marking extends to nontopical objects with particular semantic features. In Section 10.1.1 of this chapter, we saw that Old Hungarian agreement was triggered by third person topical objects, and was later reanalysed as definiteness marking and extended to definite third person objects, whether they are topical or not. At this stage of grammaticalisation the connection to information structure was totally lost, so that object marking became dependent on semantic features alone: objects with the relevant semantic features are obligatorily marked. Grammatical agreement in some Bantu languages appears to have developed in the same way.

We also saw that casemarking can work in a similar manner, as was already noticed by Bossong (1991). For instance, grammatical marking of Hindi topical objects spread onto animate specific objects. In Persian, DOM originated as a marker of topical objects denoting speech act participants, but later spread onto all definite objects, including nontopical ones, and even some indefinite objects.

This spreading scenario is similar to what Mithun (1991) has proposed for active/agentive casemarking patterns. Such patterns have a semantic basis, but the initial semantic motivation can be obscured by processes of grammaticalisation. In Caddo (Caddoan), case marking of the first argument of an intransitive verb is dependent on the notion of control: arguments that are in control of the event are classified as grammatical agents and receive agent marking, while arguments that are not in control are classified as grammatical patients and receive patient marking. However, verbs with the causative suffix automatically appear with the agent case, regardless of the degree of control involved in the situation. The reason is that in most cases the causative situation does presuppose that the agent (the causer) can control the process. The agent marking on causative agents starts as a tendency reflecting the frequent association between causation and control, but later generalises to the whole class of causative verbs.

Mithun (1991) further mentions that expansion may even be restricted to individual lexical items. For example, the verb 'lose' in Caddo behaves irregularly in the sense that its first argument is always encoded as a grammatical agent. In a similar manner, in languages with DOM the marking can become intrinsically connected with certain lexical items. For instance, question words can behave differently from other objects with respect to DOM if they bear features such as animacy. In Hebrew, where DOM is generally conditioned

by definiteness, the question word ‘who?’ in the object role obligatorily bears the object marker (Aissen 2003b:453). Browne (1970) and Karimi (1990:154) note that the grammatical marker *râ* in Persian must co-occur with the object question words *ki* ‘who’ and is optional with *ci* ‘what’ as well as objects modified by the interrogative *kodum* ‘which’. These elements are clearly in focus, and cannot be analysed as topical. We analyse the spreading of *râ*-marking to animate/human *ki* ‘who’ as generalisation of grammatical marking to nontopical elements. Since topical arguments are likely to be animate, the inherently animate ‘who’ in the object function has acquired grammatical marking, even though it is nontopical. For *ci* ‘what’ and NPs with *kodum* ‘which’, *râ*-marking appears if the NP refers to a specific entity. For example, (24a) can only be interpreted as a general question, while (24b) presupposes a choice between several known objects.

- (24) a. ce xord-i?  
 what eat-2Sg  
 ‘What did you eat?’
- b. ce-ra xord-i?  
 what-RA eat-2Sg  
 ‘What did you eat?’

These examples exhibit spreading of topical marking onto a specific nontopical object.

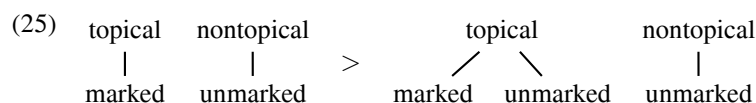
The examples discussed above clearly demonstrate that the grammaticalisation of object marking involves the familiar features of topic-worthiness: casemarking or agreement spreads onto nontopical objects that show topic-worthy features of definiteness, animacy and/or specificity. The spreading of grammatical marking to nontopics with the semantic features typical of topics is what Harris and Campbell (1995) would probably call “extension”: a change in syntax that requires generalising a rule. As they note (Harris and Campbell 1995:101), “observed extensions generalise to a natural class based on categories already relevant to the sphere in which the rule applied before it was extended”. The causal mechanism for this type of historical development is frequency (see Haspelmath 2004 and references therein on the general relevance of frequency for grammaticalisation). Topic-worthy objects are most frequent topics and therefore are most often marked at the first stage of grammaticalisation. Subsequently, the frequent association of marking with particular classes of objects changes from being preferred to being obligatory. Similar reasoning applies to individual lexical items. To account for such processes, Haspelmath suggests the **Frequency Condition on Grammaticalization**:

The more frequent a candidate for grammaticalisation is relative to other competing candidates, the more likely it is that grammaticalisation will take place. (Haspelmath 2004)

According to this condition, patterns which are more likely to be produced by speakers are therefore more likely to be “entrenched and automatized”, which ultimately leads to grammaticalisation.

### 10.2.2 *Narrowing of DOM*

In some languages only a subset of topical objects is formally marked, while nontopics must remain unmarked. We believe this is due to the historical process of “narrowing”, by which the marking becomes available only to some topics. This process is opposite to spreading, because it involves the retraction of grammatical marking. It can be schematically represented as follows:



As was shown above, in Old Hungarian and Samoyedic the original topicality-based patterns of object agreement, still found in Ob-Ugric languages, were replaced by a system where the marking of topicality was significantly reduced and restricted to third person topical objects only. We suggested that this may be due to the frequent association of secondary topicality with the third person. This process illustrates the narrowing of topical marking to a subset of topics, while all nontopical objects remain unmarked.

Narrowing processes are also observed in casemarking languages with DOM. As is well known, a number of Romance languages have differential object casemarking involving the preposition *a* (Bossong 1985 and others). As in Indo-Aryan and Iranian, it originates as an indirect object marker, but in later stages marks discourse-prominent direct objects. Further development differs across Romance languages.<sup>3</sup> A relevant example of narrowing is presented by Catalan. In Old Catalan (the period from the Middle Ages to the 19th century) *a* was attested in more contexts than those accepted by current standards: it was found on pronominals, proper nouns and human objects, although in all of these functions it was optional and correlated with topicality. This is

<sup>3</sup>Aissen (2003b) and Escandell-Vidal (2009) show that topicality-based DOM was present in Medieval Spanish. The patterns of DOM in modern Spanish have been extensively discussed in the literature, but are rather elusive, and contradictory evidence is cited in available sources; moreover, conditions on marking differ across the dialects of Spanish. We will not attempt to account for DOM in Spanish, though see Mayer (2008) for some discussion.

still observed in some varieties of the language: in Balearic Catalan casemarking appears on pronominal objects and (a subset of) lexical topical objects (Escandell-Vidal 2009). But in other dialects of modern Catalan, object marking became more restricted. In Central Catalan, the colloquial variety spoken in Barcelona and adjacent areas, DOM is based on humanness/animacy (Næss 2004). In modern Standard Catalan (the literary variety created by the normalisation process that took place at the beginning of the 20th century) only pronominal objects take the preposition *a*, and in this function it is obligatory (Escandell-Vidal 2009, Aissen 2003a,b). This is independent of information structure requirements, as the referential status of the object unambiguously determines its marking. We can conclude, then, that these varieties of Catalan demonstrate regression of object marking. At an early stage, *a* marked topical objects. But in Central Catalan and modern Standard Catalan casemarking was narrowed to human and pronominal objects, respectively, and finally lost its connection to information structure.

These examples show that, like spreading, narrowing usually involves the most typical members of the set. Since objects ranked high on prominence hierarchies are frequent topics, grammatical marking can become restricted to them. The narrowing scenario also demonstrates the non-arbitrary relation between grammaticalisation and frequency: in the words of Du Bois (1987), “grammars code best what speakers do most”. Both spreading and narrowing are based on the idea that topicality frequently correlates with topic-worthiness. The difference is that in the spreading scenario topical marking expands onto objects with frequent features of topics, while in the narrowing scenario it is retracted from topical objects with infrequent features of topics. In both cases, however, the information structure based system of DOM is lost (partially or completely) and the role of referential semantics increases.

Like spreading, the retraction of topical marking may apply in syntactic contexts determined by individual lexical items. For instance, Mahootian (1997) notes that the objects of certain Persian verbs never take *râ*, even if they are very high on the Prominence Scales. The object of the complex predicate *dombal-e gâštæn* ‘look for’ must remain unmarked even if it is animate and definite.

- (26) *dombal-e madær-æm / \*madær-æm-o mi-gærd-æm*  
 after-Ez mother-1Sg mother-1Sg-RA Dur-turn-1Sg  
 ‘I’m looking for my mother.’ (Mahootian 1997:199)

This may be due to the fact that the verb ‘look for’ often requires an indefinite nontopical object - at least in the meaning ‘look for something new’. Therefore it frequently cooccurs with unmarked objects. Then, by analogy, topical marking is retracted from all objects of this verb, even when they are definite.

Note that the standard markedness account also predicts that variations in the cross-linguistic patterns of DOM have to do with the extent to which prominence features are relevant. In Aissen's (2003b) Optimality-Theoretic analysis, historical changes are assumed to occur due to re-ranking of constraints. Differences in constraint ranking in various historical periods are characterised as the demotion or promotion of the economy constraint \*STRUC<sub>C</sub> with respect to markedness constraints. But independently of the direction of this historical change, if an object at some rank of the prominence hierarchy is formally marked, then higher ranked objects are also marked. In Aissen's own words, "the generalisation that more prominent direct objects are always more likely to be casemarked than ones of lower prominence should hold at all stages" (Aissen 2003b:471).

However, we have seen some patterns which would be difficult to account for using Prominence Scales. In Tundra Nenets, Nganasan and Selkup, object agreement is only observed with third person topics, while first and second person objects never trigger agreement even if they are topical. As demonstrated in Section 10.1.1 of this chapter, Hungarian shows further historical development, but again only third person objects participate in agreement. Yet first and second person pronouns outrank third person NPs on the definiteness hierarchy (Silverstein 1976, Aissen 1999), so they would be expected to receive more marking than the latter. There are other instances where object marking is restricted to the lower segments on the scales. For instance, Nganasan lexical objects take the accusative case while personal pronouns lack it (Tereshchenko 1979). The verb in Waris (Trans-New Guinea) agrees with animate objects if they are nouns or third person pronouns, but does not agree with first and second person pronominal objects (Brown 1988). Siewierska (2004:150) identifies a number of other languages (including Sursunga (Oceanic), Nanggu (Papuan), Waura and Parecis (Carib), and more) where object agreement targets only third person objects, while first and second person objects do not agree.

Such cases have been referred to as "inverse differential case marking" (Jäger 2003:253). The predictions that Aissen's (2003b) analysis makes are not borne out in these languages, and they are generally problematic for any markedness account, as noted by Bickel (2008:204). Optimality-style analyses based on Prominence Scales cannot provide a systematic explanation for inverse differential casemarking and, indeed, any idiosyncratic facts.

In contrast, we maintain that the direction of change cannot be predicted with certainty. Our approach is based on the idea that linguistic constructions result from individual historical processes conditioned by various factors which may be in conflict, and need not conform to markedness principles. Cross-linguistic patterns of DOM arise through the interaction of general grammaticalisation tendencies and language-particular constraints on individ-

ual constructions, as is consistent with the theory of grammatical archetypes proposed in Ackerman and Webelhuth (1998). These constraints may have different diachronic sources of a phonological or morphological nature, and these give rise to a certain amount of non-predictable variation.

For instance, we have argued that Samoyedic and Hungarian languages have grammaticalised the tendency for the first and second person topics to be preferred primary topics rather than secondary topics and therefore not to be expressed as (agreeing) objects. The tendency has grammatical consequences only in some languages: other languages grammaticalise the tendency for first and second person elements, as opposed to non-speech act participants, to be highly topical. The explanation for the Nganasan casemarking restricted to lexical objects is different. Filimonova (2005), citing personal communication with Eugene Helimsky, states that in older varieties of Nganasan object pronouns were accusatively marked, and that the loss of the accusative is due to the analogical levelling of the case paradigm. It is not clear how a markedness analysis would account for these differences.

### 10.3 Towards a typology

The diachronic scenarios outlined in the previous sections give us the following three types of languages with DOM:

**Type 1** Languages where DOM is regulated solely by information structure; correlations with semantic features are only tendencies (no spreading or narrowing).

**Type 2** Languages where DOM is regulated solely by semantic features; correlations with information structure are only tendencies (loss of connection to information structure role via narrowing or spreading).

**Type 3** Languages where DOM is regulated both by information structure and semantics:<sup>4</sup>

- (a) Languages where DOM applies to topical objects and nontopical objects with certain semantic features (spreading to arguments with topic-worthy features, while retaining connection to information structure role).
- (b) Languages where DOM applies to topical objects only if they have certain semantic features (narrowing to arguments with topic-worthy features bearing the appropriate information structure role).

---

<sup>4</sup>Presumably this is what is implied in the following quotation from Aissen (2003b:460–461): “even in zones where DOM is optional, it is often the case that the probability of casemarking varies depending on the particular properties of the object.”

As we have seen, analyses that depend purely on semantic factors cannot fully account for the patterns of DOM in languages of the first and third type.

The first type is represented by Ostyak and Vogul. There are no semantic restrictions on DOM in these languages: all semantic types of objects can trigger secondary agreement. The only exception is provided by nonreferential objects, but this follows from an independent requirement for topics to be referential. Such languages exhibit what de Hoop and Malchukov (2007) calls the fluid type of differential marking. In the fluid type the same NP can receive alternative encoding depending on the context, with a concomitant pragmatic difference.

In the second type, which de Hoop and Malchukov (2007) call “split”, different classes of NPs induce different marking, so that the marking patterns depend entirely on inherent properties of the objects. This type is exemplified by modern Hungarian. Other examples include Hebrew accusative marking on definite objects (Aissen 2003b), object agreement restricted to first person singular objects in Imbabura Quechua (Quechuan) (Cole 1982) and casemarking on first and second person objects in Yidiny (Pama-Nyungan) (Comrie 1979). More complicated patterns involving multiple features are also found. For example, Palauan object agreement appears (in the perfective aspect) if the object is either human or singular specific (Woolford 2000). In Komi-Zyrjan (Uralic), accusative case occurs either on animate or on definite objects (definiteness is marked by the third person possessive affix), while inanimate indefinite objects remain in the nominative (Toldova and Serdobolskaya 2008). We do not claim that all of these instances of semantically-based DOM originate as information structure marking: it is possible that object marking was always triggered by semantic features. The historical scenarios outline above cannot be excluded; however, since expansion and narrowing involve similar features, in the absence of clear historical evidence it is often impossible to tell which processes have taken place in languages where the connection to information structure is totally lost.

The third type seems to be the most common. In these languages, DOM is generally motivated by referential semantics, but some semantic classes allow apparent “optionality”: objects with the same semantic features are either marked or unmarked, depending on their information structure role. This type can be labelled “mixed”, i.e. it is both fluid and split. As we have seen, examples of such languages are Hindi, Chatino, Khalkha Mongolian, Tundra Nenets, Dolakha Newar and Tigre.

Subtype 3a is represented by some Bantu languages. KiSwahili object agreement is obligatory for animate objects, and optional otherwise. For inanimate objects, agreement marks discourse salience (topicality) (Seidl and Dimitriadis 1997). In the Imithupi dialect of Makua, the object marker on the verb is optional for objects of non-human classes and obligatory for human objects,



even if the overt object is nontopical (Morimoto 2002). This is demonstrated in example (27), in which the object is clearly focused but the object marker on the verb cannot be omitted.

- (27) Aráárima a-\*(n-)-líh-íre mpáni?  
 Araarima Subj-Obj-feed-Tense.Asp who  
 ‘Who did Araarima feed?’ (Morimoto 2002:(4a,b))

Given the historical scenario we have outlined, we can suspect that topical marking spread onto human objects in Imithupi, and all animate objects in KiSwahili.

In two related Austronesian languages, Selayarese and Makassarese, the verb shows object agreement with topical definite objects (Finer 1997). Focused definite objects and indefinite objects do not trigger agreement.<sup>5</sup> The restriction of topicality marking to definite topics reflects the high frequency of definite topics in discourse, although indefinite topics are in principle possible (see Chapter 3). Finer (1997) does not explain how indefinite topics behave, but in any case agreement is impossible with indefinite objects. Similarly, Aissen (2003a) shows that Sinhala casemarking is optional on animate objects, but impossible on inanimates. On our account, the Sinhala pattern may have resulted from narrowing of topicality marking: only animate topics come to be casemarked, while inanimate topics remain unmarked. Sinhala, Selayarese and Makassarese all belong to Subtype 3b, as does Tigre, where casemarking is restricted to definite topical objects.

Roughly the same categories appear to be present in languages where agreement and casemarking are not restricted to the grammatical function of object. For instance, topicality marking in Tariana does not seem to depend on any semantic restrictions, so this language would instantiate Type 1. The split Type 2 is represented by the extinct language Tangut (Tibeto-Burman), where agreement on the verb is triggered by first and second person arguments alone; the agreement controller can correspond to a variety of grammatical functions including the subject, the object, the possessor of the subject and the possessor of the object (Kepping 1979).<sup>6</sup>

<sup>5</sup>Focus and nonfocus/topical objects differ in their linear position in Selayarese.

<sup>6</sup>Kepping (1979) does not provide glosses for these examples; we reproduce Comrie’s (2003) glosses for these examples, including for the morpheme glossed only as ‘?’.

- (28) a.  $\text{thin}^2 \text{siwo}^2 \text{t\acute{s}ia}^1 \quad \text{nd\acute{z}iwo}^2 \text{mi}^1 \text{nd\acute{z}iu}^1\text{-nga}^2$   
 why virtuous people Neg love-1Sg  
 ‘Why do I not love virtuous people?’ (Kepping 1979:268)
- b.  $\text{nga}^2.\text{in}^1 \text{\acute{s}a}^1\text{men}^1 \text{kwje}^1 \text{ndi}^2\text{-khion-nga}^1$   
 to.me \acute{s}ramana fruit give-?-1Sg  
 ‘Give the fruit of the \acute{s}ramana to me.’ (Kepping 1979:269)
- c.  $\text{nd\acute{z}iwo}^2 \text{ngi}^2 \text{nga}^2.\text{in}^1 \text{lda}^2 \text{ki}^1\text{-\acute{z}won}^2\text{-nga}^2$   
 someone my hand grasped-?-1Sg  
 ‘Someone grasped my hand.’ (Kepping 1979:270)

Thus, in Tangut only first and second person arguments trigger agreement, but agreement is “trigger-happy” in terms of grammatical functions. The mixed Type 3 includes Persian and a number of other languages.

#### 10.4 Conclusion

Though we have not provided an explicit formalisation of the processes of linguistic change we have proposed in this chapter, we agree with Vincent (1997) that LFG provides a fruitful theoretical setting for the exploration of these processes. As Vincent notes, LFG is well equipped to handle the lexical basis and lexically specified exceptions to the processes we describe. Our precise formal model of the relations between c-structure, f-structure, semantic structure, and information structure provides a good basis for observing interactions among these levels and the effect of features at different levels on the formulation and reinterpretation of constraints involving the levels.

Our theory of the historical genesis of DOM is not complete: we are aware that in many languages other factors play a role in DOM, for instance, volitionality and control on the part of the agent participant and the factors related to the structure of event (aspect, telicity, and incremental theme). The relationship of these factors to topicality and DOM is yet to be explored.

## *Conclusion*

We have explored the effect of information structure on grammatical marking, presenting evidence from languages that treat topics specially in terms of grammatical marking. Topicality is a relational property of a referent, determined by the speaker's assessment of its relative saliency, and cannot be "measured" in terms of inherent semantic features such as animacy: topical referents are what propositions are construed to be about. Crucial to our analysis is the possibility for more than one sentence element to be topical. We distinguish between the **primary topic** and the **secondary topic**; both are topics, but the primary topic is more pragmatically prominent. Although there is no unique alignment between information structure roles and grammatical functions, there are important cross-linguistic tendencies in the grammatical expression of primary and secondary topics: in particular, we have argued that while subjects are prototypical/canonical primary topics, objects tend to be associated with secondary topics.

In the simplest cases of apparently "optional" case- and adpositional marking and agreement, the factor determining the presence of marking is whether a sentence element is topical. In some languages, casemarking and agreement mark the topical status of any grammatical function, subjects as well as non-subjects. Other languages grammatically mark topicality for a range of non-subject NPs. Restrictions on marking in these cases have often been treated in syntactic terms, but we believe that some of these apparent syntactic restrictions may be better thought of as a consequence of independent constraints on how topics can be syntactically realised.

Some languages place additional syntactic requirements on the grammatical marking of nonsubject topics, restricting marking to object-like grammatical functions. This gives rise to **differential object marking** (DOM): casemarking and agreement patterns in many languages with DOM distinguish topical

objects, which are grammatically marked, from nontopical, grammatically unmarked objects. In emphasising the role of information structure in DOM, our analysis accounts for cases of apparent optionality which have not been fully addressed by semantically-driven proposals. In particular, we account for patterns of DOM in languages with “local” systems of object marking with no obvious semantic motivation, showing that such patterns are determined by information structure role. Optimality-theoretic approaches which rely on referential semantic features of the object handle such cases in terms of reranking of constraints, leading to unexplained optionality of marking. On our analysis, “true” optionality does not arise: objects with the same semantic features are either grammatically marked or unmarked, depending on their information structure role.

Our approach provides a unified account of topical marking that accounts not only for DOM, but also for languages where marking of a variety of grammatical dependents depends on topicality: it accounts unproblematically for languages like Persian, Itelmen and Tariana, where topic marking is not restricted to objects, as well as languages exhibiting the classic patterns of DOM. It is not clear how previous analyses of DOM can easily be extended to account for these languages.

Altogether, our analysis provides a motivation for DOM that is different from the claims of much previous research. Most work on DOM assumes that object marking originates from the need to differentiate the object from the subject. However, we claim that DOM actually marks similarities rather than differences between subjects (canonical topics) and topical objects: topics tend to bear grammatical marking, no matter what their grammatical function. Thus, our analysis does not relate the formal markedness of objects with their functional markedness, at least if the latter is assessed in terms of frequency or typicality. Instead, it highlights the **coding or indexing** function of marking as an indicator of topicality. Our approach stands in opposition to the common view that objects are prototypically aligned with the focus function: we have argued that the *SUBJ/topic*, *OBJ/(secondary) topic* alignment is equally likely, where both core arguments are topical. In support of this view, we have discussed evidence that shows that topical objects are at least as frequent in discourse as focused objects, and in this sense cannot be considered functionally marked. In this, our analysis is in line with recent work emphasising the coding function of marking in the related phenomenon of Differential Subject Marking: many of the contributors to a recent collection of papers on Differential Subject Marking (de Hoop and de Swart 2009) argue that not all Differential Subject Marking effects can be attributed to the disambiguating function of grammatical marking, and this is exactly what we have found for DOM.

While DOM patterns in some languages are defined in purely information structural terms, in other languages they depend on semantic features such as

animacy, definiteness, and the like, or on a combination of information structure and semantic features. We have suggested that marking based purely on information structure roles is historically primary, at least in some languages, and that the importance of semantic features emerges as a result of different directions of grammaticalisation of topic marking. Grammatical marking is either extended to nontopical objects characterised by certain semantic features, or restricted to a subset of topical arguments. The role of semantic features in DOM is thus explained by historical patterns of grammaticalisation: typical semantic features of topics come to be relevant for grammatical marking.

Another crucial difference between our work and many previous analyses of DOM is that we do not discuss only grammatical (morphological) marking of objects, but pay special attention to their behavioural syntax. Typologically based work does not usually address the syntax of objects, while most generative research concentrates on positional differences. Our analysis does not define DOM in terms of object position because we do not assume that syntactic roles are defined configurationally: following the standard LFG view, we take grammatical functions to be primitives which are not defined in terms of their syntactic position.

In our investigation of the grammatical behaviour of grammatically marked and unmarked objects, we found that languages differ: in some languages they are both **primary objects**, while in other languages they bear different object-like functions. In languages like Ostyak, Khalkha Mongolian and Chatino, grammatical marking of objects may seem to depend on information structure: topical objects are marked, while nontopical objects are unmarked. However, closer examination reveals that, in fact, marking patterns in these languages are defined in completely syntactic terms, just as in English or Latin. The distinguishing characteristic of these languages is the obligatory linkage between grammatical functions and information structure: primary objects are always topical, while secondary or restricted objects are nontopical. This means that in some cases grammatical structure may arise diachronically under pressure from information structure constraints. The need to distinguish two types of information structuring (with and without a topical object) has led to grammatical differences that go beyond patterns of agreement or casemarking.

This means that theories of argument mapping, relating semantic roles to grammatical functions, must take into account not only the semantic role of the argument of a predicate but also its information structure role in determining its grammatical function. This is in line with work by Morimoto (2009:212), who proposes to distinguish topic-prominent languages in which the topic is linked to subject from subject-prominent languages in which the thematically highest argument is linked to subject. We argue that the same is observed with objects: in some languages, linking to the grammatical function of object

is strongly influenced by information structure role, while in other languages semantic role is the sole determinant of linking patterns.

We have also clarified the position of information structure within the grammatical architecture of LFG, and shown how constraints from various levels of linguistic structure can combine to specify and determine information structure role. LFG distinguishes grammatical marking, grammatical function, and information structure role, which has been of crucial importance in formulating our theory. Abstract grammatical functions are not assumed to correlate one-to-one with case or agreement morphology. They are not defined in terms of their phrase structural position in the sentence or in terms of morphological properties.

In sum, our approach is different from much previous work, in which case-marking and agreement patterns are treated in terms of purely syntactic factors (sometimes in combination with semantics), by reference to syntactic roles like subject and object. This works unproblematically for languages where agreement patterns do not depend on information structure role. Our work shows that these processes in fact can make reference to other levels of structure, but that careful examination is needed to determine whether they relate directly to information structure, or only indirectly, by virtue of a tight alignment between informational roles and grammatical functions. Thus, we see that although agreement and casemarking are syntactically constrained, different languages can exploit the syntax-information structure interface in different ways to determine patterns of grammatical marking. Examining these patterns illuminates our understanding of the syntax-information structure interface.

## References

- Ackerman, Farrell and Webelhuth, Gert 1998. *A Theory of Predicates*. Stanford, CA: CSLI Publications
- Ahmed, Tafseer 2006. 'Spatial, temporal and structural uses of Urdu *ko*.' URL <http://csli-publications.stanford.edu/LFG/11/1fg06ahmed.pdf>. Workshop on case and aspect in South Asian languages, LFG06, Konstanz
- Aikhenvald, Alexandra Y. 2003. *A grammar of Tariana, from Northwest Amazonia*. Cambridge, UK: Cambridge University Press
- Aissen, Judith 1999. 'Markedness and subject choice in optimality theory.' *Natural Language and Linguistic Theory* 17(4):673–711. Also in Legendre et al. (2001)
- Aissen, Judith 2003a. 'Differential coding, partial blocking, and bidirectional OT.' In Nowak, Pawel and Yoquelet, Corey (eds.) 'Proceedings of the 29th Annual Meeting of the Berkeley Linguistics Society,' University of California, Berkeley: Berkeley Linguistics Society
- Aissen, Judith 2003b. 'Differential object marking: Iconicity vs. economy.' *Natural Language and Linguistic Theory* 21(3):435–483
- Alsina, Alex 1996. *The Role of Argument Structure in Grammar: Evidence from Romance*. Stanford, CA: CSLI Publications
- Alsina, Alex 2008. 'A theory of structure-sharing: Focusing on long-distance dependencies and parasitic gaps.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2008 Conference,' URL <http://csli-publications.stanford.edu/LFG/13/1fg08.html>
- Andrews, III, Avery 1982. 'The representation of case in modern Icelandic.' In Bresnan, Joan (ed.) 'The Mental Representation of Grammatical Relations,' Cambridge, MA: The MIT Press, pp. 427–503

- Andrews, III, Avery 1985. 'The major functions of the noun phrase.' In Shopen, Tim (ed.) 'Language Typology and Syntactic Description,' Cambridge, UK: Cambridge University Press
- Ariel, Mira 1988. 'Referring and accessibility.' *Journal of Linguistics* 24:67–87
- Asher, R. E. and Kumari, T. C. 1997. *Malayalam*. London and New York: Routledge
- Asudeh, Ash 2004. *Resumption as Resource Management*. Ph.D. thesis, Stanford University
- Asudeh, Ash 2005. 'Relational nouns, pronouns, and resumption.' *Linguistics and Philosophy* 28(4):375–446
- Asudeh, Ash 2006. 'Direct compositionality and the architecture of LFG.' In Butt, Miriam, Dalrymple, Mary, and King, Tracy Holloway (eds.) 'Intelligent Linguistic Architectures: Variations on themes by Ronald M. Kaplan,' Stanford, CA: CSLI Publications
- Austin, Peter and Bresnan, Joan 1996. 'Non-configurationality in Australian Aboriginal languages.' *Natural Language and Linguistic Theory* 14(2):215–268
- Bárczi, Géza 1980. *A magyar nyelv múltja és jelene [The past and present of the Hungarian language]*. Budapest: Gondolat
- Barðdal, Jóhanna 2001. 'Case in Icelandic: A synchronic, diachronic, and comparative approach.' In 'Lundastudier i Nordisk Språkvitenskap,' Department of Scandinavian Languages, Lund University, volume 57
- Beames, John 1966 [1872-79]. *A Comparative Grammar of the Modern Aryan Languages of India*. Delhi: Munshiram Manoharlal
- Beaver, David I. and Clark, Brady Z. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Chichester: Wiley-Blackwell
- Beck, David 2006. 'Control of agreement in multi-object constructions in Upper Necaxa Totonac.' In Fujimori, Atsushi and Silva, Maria Amelia Reis (eds.) 'Proceedings of the 11th Workshop on Structure and Constituency in the Languages of the Americas,' Vancouver, pp. 1–11
- Beck, David 2007. 'Voice and agreement in multi-object constructions in Upper Necaxa Totonac.' In 'Memorias de IX Encuentro Internacional de Lingüística en el Noroeste,' Hermosillo



- Beck, David 2008. 'Sorting out grammatical relations in multi-object constructions in Upper Necaxa Totonac.' On-line festschrift for Bernard Comrie
- Berman, Judith 1999. 'Does German satisfy the Subject Condition?' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG99 Conference,' URL <http://csli-publications.stanford.edu/LFG/4/lfg99.html>
- Bhatt, Rajesh and Anagnostopoulou, Elena 1996. 'Object shift and specificity: Evidence from ko-phrases in Hindi.' In 'Papers from the Thirty-Second Regional Meeting of the Chicago Linguistic Society,' University of Chicago: Chicago Linguistic Society, pp. 11–22
- Bickel, Balthasar 2008. 'On the scope of the referential hierarchy in the typology of grammatical relations.' In Corbett, Greville and Noonan, Michael (eds.) 'Case and grammatical relations,' Amsterdam: John Benjamins, pp. 191–210
- Bickel, Balthasar, Bisang, Walter, and Yadava, Yogendra P. 1999. 'Face vs. empathy: the social foundation of Maithili verb agreement.' *Linguistics* 37(3):481–518
- Blake, Barry J. 2001. *Case*. Cambridge, UK: Cambridge University Press, second edition
- Bobaljik, Jonathan David and Wurmbrand, Susi 2002. 'Notes on agreement in Itelmen.' *Linguistic Discovery* 1(1). URL <http://linguistic-discovery.dartmouth.edu>
- Bokamba, Eyamba G. 1981. 'Aspects of Bantu syntax.' Unpublished manuscript, cited in Bresnan and Mchombo (1987)
- Börjars, Kersti, Chisarik, Erika, and Payne, John 1999. 'On the justification for functional categories in LFG.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG99 Conference,' URL <http://csli-publications.stanford.edu/LFG/4/lfg99.html>
- Börjars, Kersti and Vincent, Nigel 2005. 'Position vs. function in Scandinavian Presentational Constructions.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2005 Conference,' URL <http://csli-publications.stanford.edu/LFG/10/lfg05.html>
- Börjars, Kersti and Vincent, Nigel 2008. 'Objects and OBJ.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2008 Conference,' URL <http://csli-publications.stanford.edu/LFG/13/lfg08.html>

- Bos, Johan, Mastenbroek, Elsbeth, McGlashan, Scott, Millies, Sebastian, and Pinkal, Manfred 1994. 'A compositional DRS-based formalism for NLP-applications.' In 'International Workshop on Computational Semantics,' Tilburg. Also published as *Verbmobil Report 59*, Universität des Saarlandes, Saarbrücken, Germany
- Bossong, Georg 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr
- Bossong, Georg 1989. 'Morphemic marking of topic and focus.' In Kefer, Michel and van der Auwera, Johan (eds.) 'Belgian Journal of Linguistics 4: Universals of Language,' Amsterdam: John Benjamins
- Bossong, Georg 1991. 'Differential object marking in Romance and beyond.' In Wanner, Dieter and Kibbee, Douglas A. (eds.) 'New Analyses in Romance Linguistics,' Amsterdam: John Benjamins, pp. 143–167
- Bowern, Claire 2004. *Bardi Verb Morphology in Historical Perspective*. Ph.D. thesis, Harvard University
- Bresnan, Joan 1980. 'Polyadicity.' In Hoekstra, Teun, van der Hulst, Harry, and Moortgat, Michael (eds.) 'Lexical Grammar,' Dordrecht: Foris Publications, pp. 97–121. Republished in revised form in Bresnan (1982)
- Bresnan, Joan (ed.) 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: The MIT Press
- Bresnan, Joan 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers
- Bresnan, Joan and Kanerva, Jonni M. 1989. 'Locative inversion in Chicheŵa: A case study of factorization in grammar.' *Linguistic Inquiry* 20(1):1–50. Reprinted in Stowell et al. (1992)
- Bresnan, Joan and Mchombo, Sam A. 1987. 'Topic, pronoun, and agreement in Chicheŵa.' *Language* 63(4):741–782
- Bresnan, Joan and Moshi, Lioba 1990. 'Object asymmetries in comparative Bantu syntax.' *Linguistic Inquiry* 21(2):147–186
- Bresnan, Joan and Zaenen, Annie 1990. 'Deep unaccusativity in LFG.' In Dziwirek, Katarzyna, Farrell, Patrick, and Mejías-Bikandi, Errapel (eds.) 'Grammatical Relations: A Cross-Theoretical Perspective,' Stanford, CA: CSLI Publications, pp. 45–57
- Brown, Robert 1988. 'Waris case system and verb classification.' *Language and Linguistics in Melanesia* 19:37–80

- Browne, Wayles 1970. 'More on definiteness markers: Interrogatives in Persian.' *Linguistic Inquiry* 1(3):359–363
- Büring, Daniel 2007. 'Semantics, intonation, and information structure.' In Ramchand, Gillian and Reiss, Charles (eds.) 'The Oxford Handbook of Linguistic Interfaces,' Oxford: Oxford University Press, pp. 445–474
- Butt, Miriam 1993. 'Object specificity and agreement in Hindi/Urdu.' In Beals, K. et al. (eds.) 'Papers from the Twenty-Ninth Regional Meeting of the Chicago Linguistic Society,' University of Chicago: Chicago Linguistic Society, pp. 89–103
- Butt, Miriam 1995. *The Structure of Complex Predicates in Urdu*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Revised and corrected version of 1993 Stanford University dissertation.
- Butt, Miriam 2008a. 'From spatial expression to core case marker: Ergative and dative/accusative.' URL <http://ling.uni-konstanz.de/pages/home/butt/>. MS, University of Konstanz
- Butt, Miriam 2008b. *Theories of Case*. Cambridge, UK: Cambridge University Press
- Butt, Miriam, Ahmed, Tafseer, and Poudel, Tikaram 2008. 'Development of case in South Asian languages.' URL <http://ling.uni-konstanz.de/pages/home/tafseer/dgfs08-hnd.pdf>. Paper presented at DGfS Workshop *Sprachwandelvergleich* Bamberg, February 28th, 2008
- Butt, Miriam, Grimm, Scott, and Ahmed, Tafseer 2006. 'Dative subjects.' URL <http://ling.uni-konstanz.de/pages/home/butt/nijmegen-hnd.pdf>. Paper presented at NWO/DFG Workshop on Optimal Sentence Processing
- Butt, Miriam and King, Tracy Holloway 1996. 'Structural topic and focus without movement.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG96 Conference,' URL <http://csli-publications.stanford.edu/LFG/1/lfg1.html>
- Butt, Miriam and King, Tracy Holloway 1999. 'The status of case.' In Dayal, Veneeta and Mahajan, Anoop (eds.) 'Clause Structure in South Asian Languages,' Dordrecht: Kluwer Academic Publishers, pp. 153–198
- Butt, Miriam and King, Tracy Holloway 2000. 'Null elements in discourse structure.' In Subbarao, K. V. (ed.) 'Papers from the NULLS Seminar,' Delhi: Motilal Banarsidass

- Butt, Miriam and King, Tracy Holloway 2003. 'Case systems: Beyond structural distinctions.' In Brandner, Ellen and Zinsmeister, Heike (eds.) 'New Perspectives on Case Theory,' Stanford, CA: CSLI Publications, pp. 53–87
- Butt, Miriam, Niño, María-Eugenia, and Segond, Frédérique 1996. 'Multilingual processing of auxiliaries in LFG.' In Gibbon, D. (ed.) 'Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference,' Mouton de Gruyter, pp. 111–122. Universität Bielefeld, 7 - 9 October 1996
- Carleton, Troi and Waksler, Rachelle 2000. 'Pronominal markers in Zenzontepec Chatino.' *International Journal of American Linguistics* 66(3):381–395
- Carleton, Troi and Waksler, Rachelle 2002. 'Marking focus in Chatino.' *Word* 53(2):157–171
- Çetinoğlu, Özlem and Butt, Miriam 2008. 'Turkish non-canonical objects.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2008 Conference,' URL <http://csli-publications.stanford.edu/LFG/13/lfg08.html>
- Choi, Hye-Won 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Revised and corrected version of 1996 Stanford University dissertation.
- Chomsky, Noam 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press
- Chomsky, Noam and Lasnik, Howard 1993. 'The theory of principles and parameters.' In Jacobs, Joachim, von Stechow, Arnim, Sternefeld, Wolfgang, and Vennemann, Theo (eds.) 'Syntax: An International Handbook of Contemporary Research,' Cambridge, MA: Walter de Gruyter
- Cole, Peter 1982. *Imbabura Quechua*. Amsterdam: North Holland
- Comrie, Bernard 1977. 'Subjects and direct objects in Uralic languages: A functional explanation of case-marking systems.' *Études Finno-Ougriennes* 12:5–17
- Comrie, Bernard 1979. 'Definite and animate direct objects: A natural class.' *Linguistica Silesiana* III:13–21
- Comrie, Bernard 1989. *Language universals and linguistic typology*. Chicago: University of Chicago Press, second edition

- Comrie, Bernard 2003. 'When agreement gets trigger-happy.' *Transactions of the Philological Society* 101(2):313–337
- Cook, Phillipa and Payne, John 2006. 'Information structure and scope in German.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2006 Conference,' URL <http://csli-publications.stanford.edu/LFG/11/lfg06.html>
- Corbett, Greville G. 2006. *Agreement*. Cambridge, UK: Cambridge University Press
- Creissels, Denis 2008. 'Remarks on split intransitivity and fluid intransitivity.' In Bonami, Olivier and Hofherr, Patricia Cabredo (eds.) 'Empirical issues in syntax and semantics,' CSSP, volume 7, pp. 139–168. URL [http://www.cssp.cnrs.fr/eiss7/index\\_en.html](http://www.cssp.cnrs.fr/eiss7/index_en.html)
- Croft, William 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press
- Croft, William A. 1988. 'Agreement vs. case marking and direct objects.' In Barlow, Michael and Ferguson, Charles A. (eds.) 'Agreement in Natural Language: Approaches, Theories, Descriptions,' Stanford, CA: CSLI Publications, pp. 159–179
- Croft, William A. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press
- Culicover, Peter W. and Jackendoff, Ray 2005. *Simpler Syntax*. Oxford University Press
- Culy, Christopher D. 1994. 'Extraposition in English and grammatical theory.' MS, University of Iowa
- Culy, Christopher D. 2000. 'An incorporated topic marker in Takelma.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2000 Conference,' URL <http://csli-publications.stanford.edu/LFG/5/lfg00.html>
- Cummins, Sarah 2000. 'The unaccusative hypothesis and the impersonal construction in French.' *Canadian Journal of Linguistics* 45:225–251
- Dabir-Moghaddam, Mohammad 1992. 'On the (in)dependence of syntax and pragmatics: Evidence from the postposition *-rā* in Persian.' In Stein, Dieter (ed.) 'Cooperating with written texts: the pragmatics and comprehension of written texts,' Berlin: Mouton de Gruyter, pp. 549–573

- Dahl, Östen and Fraurud, K. 1996. 'Animacy in grammar and discourse.' In Fretheim, Thorstein and Gundel, Jeanette K. (eds.) 'Reference and referent accessibility,' Amsterdam: John Benjamins, pp. 47–64.
- Dahlstrom, Amy 2009. 'OBJ $\theta$  without OBJ: a typology of Meskwaki objects.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2009 Conference,' URL <http://csli-publications.stanford.edu/LFG/14/lfg09.html>
- Dalrymple, Mary (ed.) 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: The MIT Press
- Dalrymple, Mary 2001. *Lexical Functional Grammar*, volume 34 of *Syntax and Semantics*. New York, NY: Academic Press
- Dalrymple, Mary, Kaplan, Ronald M., Maxwell, III, John T., and Zaenen, Annie (eds.) 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications
- Dalrymple, Mary, Lamping, John, Pereira, Fernando C. N., and Saraswat, Vijay A. 1997. 'Quantifiers, anaphora, and intensionality.' *Journal of Logic, Language, and Information* 6(3):219–273. Reprinted in Dalrymple (1999)
- Dalrymple, Mary, Lamping, John, Pereira, Fernando C. N., and Saraswat, Vijay A. 1999. 'Overview and introduction.' In Dalrymple, Mary (ed.) 'Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach,' Cambridge, MA: The MIT Press
- Dalrymple, Mary and Nikolaeva, Irina 2005. 'Nonsubject agreement and discourse roles.' In 'Oxford Working Papers in Linguistics, Philology, and Phonetics,' University of Oxford: Centre for Linguistics and Philology, pp. 73–94
- Danon, Gabi 2006. 'Caseless nominals and the projection of DP.' *Natural Language and Linguistic Theory* 24(4):977–1008
- Dayal, Veneeta 2003. 'Bare nominals: Non-specific and contrastive readings under scrambling.' In Karimi, Simin (ed.) 'Word order and scrambling,' Oxford: Blackwell Publishers
- de Hoop, Helen 1992. *Case configuration and noun phrase interpretation*. Ph.D. thesis, University of Groningen. Reprinted by Garland Press, New York, 1996
- de Hoop, Helen 2008. 'Against affectedness.' In 'Cognitive and Functional Perspectives on Dynamic Tendencies in Languages,' Tartu: University of Tartu and ECLA, pp. 212–213

- de Hoop, Helen and de Swart, Peter (eds.) 2009. *Differential Subject Marking*. Dordrecht: Springer
- de Hoop, Helen and Lamers, Monique 2006. 'Incremental distinguishability of subject and object.' In Kulikov, Leonid, Malchukov, Andrej, and de Swart, Peter (eds.) 'Case, valency and transitivity,' Amsterdam: John Benjamins, pp. 269–287
- de Hoop, Helen and Malchukov, Andrej 2007. 'On fluid differential case marking: A bidirectional OT approach.' *Lingua* 117:1636–1656
- de Hoop, Helen and Narasimhan, Bhuvana 2005. 'Differential case-marking in Hindi.' In Amberber, Mengistu and de Hoop, Helen (eds.) 'Competition and variation in natural languages: the case for case,' London and Amsterdam: Elsevier, pp. 321–345
- de Swart, Peter 2006. 'Case markedness.' In Kulikov, Leonid, Malchukov, Andrej, and de Swart, Peter (eds.) 'Case, valency and transitivity,' Amsterdam: John Benjamins, pp. 249–267
- de Swart, Peter 2007. *Cross-linguistic Variation in Object Marking*. Ph.D. thesis, Radboud University Nijmegen
- Deo, Ashwini 2008. 'Datives: Locations and possessors. Case syncretism in Marathi diachrony.' URL [http://ling.uni-konstanz.de/pages/home/tafseer/workshop/deo\\_case\\_2008.pdf](http://ling.uni-konstanz.de/pages/home/tafseer/workshop/deo_case_2008.pdf). Paper presented at SFB 471 Case Workshop University of Konstanz, May 8-10 2008
- Diesing, Molly 1992. *Indefinites*. Cambridge, MA: The MIT Press
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge, UK: Cambridge University Press
- Dobrovie-Sorin, Carmen 1994. *The syntax of Romanian*. Berlin: Mouton de Gruyter
- Donohue, Mark 2004. 'A voice opposition without voice morphology.' Paper presented at AFLA 2004, Berlin
- Dowty, David R. 1991. 'Thematic proto roles and argument selection.' *Language* 67(3):547–619
- Dryer, Matthew S. 1986. 'Primary objects, secondary objects, and antitativity.' *Language* 62(4):808–845
- Du Bois, John W. 1987. 'The discourse basis of ergativity.' *Language* 63(4):805–855

- É. Kiss, Katalin (ed.) 1995. *Discourse Configurational Languages*. Oxford: Oxford University Press
- É. Kiss, Katalin 2002. *The syntax of Hungarian*. Cambridge, UK: Cambridge University Press
- Enç, Mürvet 1986. 'Towards a referential analysis of temporal expressions.' *Linguistics and Philosophy* 9(4):405–426
- Enç, Mürvet 1991. 'The semantics of specificity.' *Linguistic Inquiry* 22:1–26
- Engdahl, Elisabet and Vallduví, Enric 1996. 'Information packaging in HPSG.' In Grover, Claire and Vallduví, Enric (eds.) 'Studies in HPSG,' Edinburgh: University of Edinburgh, volume 12 of *Edinburgh Working Papers in Cognitive Science*
- Erteschik-Shir, Nomi 1997. *The dynamics of focus structure*. Cambridge, UK: Cambridge University Press
- Erteschik-Shir, Nomi 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press
- Escandell-Vidal, M. Victoria 2009. 'Differential object marking and topicality: The case of Balearic Catalan.' *Studies in Language* 33(4):832–884
- Falk, Yehuda N. 2001. *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI Publications
- Falk, Yehuda N. 2006. *Subjects and Universal Grammar: An Explanatory Theory*. Cambridge, UK: Cambridge University Press
- Farkas, Donca 1995. 'Specificity and scope.' In Nash, L. and Tsoulas, G. (eds.) 'Actes du Premier Colloque Langues and Grammaire,' Paris, pp. 119–137
- Féry, Caroline and Krifka, Manfred 2008. 'Information structure: Notional distinctions, ways of expression.' In van Sterkenburg, Piet (ed.) 'Unity and diversity of languages,' Amsterdam: John Benjamins, pp. 123–136
- Filimonova, Elena 2005. 'The noun phrase hierarchy and relational marking: Problems and counterevidence.' *Linguistic Typology* 9:77–113
- Finer, Daniel L. 1997. 'Contrasting  $\bar{A}$ -dependencies in Selayarese.' *Natural Language and Linguistic Theory* 15:677–728
- Frank, Anette and Zaenen, Annie 2002. 'Tense in LFG: Syntax and morphology.' In Kamp, Hans and Reyle, Uwe (eds.) 'How we say WHEN it happens: Contributions to the theory of temporal reference in natural language,' Tübingen: Niemeyer



- Fry, John 2001. *Ellipsis and WA-Marking in Japanese Conversation*. Ph.D. thesis, Stanford University
- Gair, James W. and Paolillo, John C. 1997. *Sinhala*. München: LINCOM Europa
- Genetti, Carol 1994. *A Descriptive and Historical Account of the Dolakha Newari Dialect*. Tokyo: Institute for the Study of Languages and Cultures of Africa and Asia
- Genetti, Carol 1997. 'Object relations and dative case in Dolakha Newari.' *Studies in Language* 21(1):37–68
- Genetti, Carol 2007. *A grammar of Dolakha Newar*. Mouton Grammar Series 40. Berlin: Mouton de Gruyter
- Georgopoulos, Carol 1992. 'Another look at object agreement.' In 'Proceedings of the 22nd Conference of the North Eastern Linguistic Society (NELS22),' University of Massachusetts at Amherst, pp. 165–177
- Geurts, Bart 2002. 'Specific indefinites, presupposition and scope.' In Bäuerle, Rainer, Reyle, Uwe, and Zimmermann, Thomas Ede (eds.) 'Presuppositions in discourse,' Oxford: Elsevier. To appear
- Ghameshi, Jila 1997. 'Topics in Persian VPs.' *Lingua* 102:133–167
- Givón, Talmy 1976. 'Topic, pronoun, and grammatical agreement.' In Li, Charles (ed.) 'Subject and Topic,' New York: Academic Press, pp. 57–98
- Givón, Talmy 1979. *On Understanding Grammar*. New York: Academic Press
- Givón, Talmy 1983a. 'Introduction.' In Givón, Talmy (ed.) 'Topic Continuity in Discourse: A Quantitative Cross-Language Study,' Amsterdam: John Benjamins
- Givón, Talmy (ed.) 1983b. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins
- Givón, Talmy 1984a. 'Direct object and dative shifting: semantic and pragmatic case.' In Plank, Frans (ed.) 'Objects: Towards a theory of grammatical relations,' London: Academic Press, pp. 151–182
- Givón, Talmy 1984b. *Syntax: A Functional-Typological Introduction*, volume I. Amsterdam: John Benjamins
- Givón, Talmy 1990. *Syntax: A Functional-Typological Introduction*, volume II. Amsterdam: John Benjamins

- Givón, Talmy 2001. *Syntax*. Amsterdam: John Benjamins
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press
- Golovko, Evgenij 2009. *Aleutskij jazyk v Rossijskoj Federacii (Struktura, funkcionirovanie, kontaktnye javlenija)*. [The Aleut language in the Russian Federation: Structure, functioning, contact phenomena.]. Ph.D. thesis, Saint Petersburg Academy of Sciences
- Greenberg, Joseph H. 1977. 'Niger-Congo noun class markers: Prefixes, suffixes, both or neither.' *Studies in African Linguistics, supplement 7* pp. 97–104
- Grimes, Barbara F. 1999. *ETHNOLOGUE: Languages of the World*. SIL International. URL <http://www.sil.org/ethnologue/>
- Gundel, Jeanette K. 1988. *The role of topic and comment in linguistic theory*. New York: Garland
- Gundel, Jeanette K., Hedberg, Nancy, and Zacharski, Ron 1993. 'Cognitive status and the form of referring expressions in discourse.' *Language* 69:274–308
- Guntsetseg, Dolgor 2008. 'Differential object marking in Mongolian.' In Schäfer, F. (ed.) 'SinSpec: Working Papers of the FSB732 'Incremental Specification in Context', Universität Stuttgart, volume 1, pp. 53–69
- Haig, Geoffrey 2008. *Alignment change in Iranian languages: a Construction Grammar approach*. New York: Mouton de Gruyter
- Haiman, John and Thompson, Sandra A. 1985. *Natural Syntax: Iconicity and Erosion*. Cambridge, UK: Cambridge University Press
- Hajičová, Eva, Partee, Barbara H., and Sgall, Petr 1998. *Topic-focus articulation, tripartite structures, and semantic content*. Dordrecht: Kluwer Academic Publishers
- Harris, Alice and Campbell, Lyle 1995. *Historical syntax in cross-linguistic perspective*. Cambridge, UK: Cambridge University Press
- Haspelmath, Martin 2004. 'Explaining the ditransitive person-role constraint: A usage-based account.' *Constructions* 2/2004. URL <http://www.constructions-online.de/articles/35>

- Haspelmath, Martin 2007. 'Ditransitive alignment splits and inverse alignment.' *Functions of Language* 14(1):79–102. Special issue on ditransitives, guest edited by Anna Siewierska
- Heim, Irene 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts at Amherst
- Heine, Bernd and Kuteva, Tania 2002. *World Lexicon of Grammaticalization*. Cambridge, UK: Cambridge University Press
- Her, One-Soon 1991. 'Topic as a grammatical function in Chinese.' *Lingua* 84(1):1–23
- Hewson, John and Bubenik, Vit 2006. *From case to adposition: the development of configurational syntax in Indo-European languages*. Amsterdam: John Benjamins
- Hockett, Charles 1958. *A course in modern linguistics*. New York: Macmillan
- Honti, László 1984. *Chrestomathia ostiacica*. Budapest: Tankönyvkiadó
- Hopper, Paul J. and Thompson, Sandra A. 1980. 'Transitivity in grammar and discourse.' *Language* 56:251–299
- Huang, Chu-Ren 1989. 'Subcategorized TOPICs in Mandarin Chinese.' URL <http://cwn.ling.sinica.edu.tw/churen/SubcategorizedTOPICsinMandarinChinese.pdf>. MS, Academia Sinica
- Hudson, Richard 1992. 'So-called 'double objects' and grammatical relations.' *Language* 68(2):251–276
- Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press
- Jackendoff, Ray S. 1983. *Semantics and Cognition*. Cambridge, MA: The MIT Press
- Jackendoff, Ray S. 1990. *Semantic Structures*. Cambridge, MA: The MIT Press
- Jaeger, T. Florian 2004. 'Topicality and superiority in Bulgarian wh-questions.' In Arnaudova, Olga, Browne, Wayles, Rivero, Maria Luisa, and Stojanovic, Danijela (eds.) 'Proceedings of FASL-12: The Ottawa Meeting 2003,' pp. 207–228

- Jaeger, T. Florian and Gerassimova, Veronica A. 2002. 'Bulgarian word order and the role of the direct object clitic in LFG.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2002 Conference,' URL <http://csli-publications.stanford.edu/LFG/7/lfg2.html>
- Jäger, Gerhard 2003. 'Learning constraint subhierarchies: the Bidirectional Gradual Learning Algorithm.' In Blutner, Reinhard and Zeevat, Henk (eds.) 'Optimality theory and pragmatics,' Basingstoke: Palgrave Macmillan, pp. 251–298
- Jake, Janice L. 1980. 'Object verb agreement in Tigre.' *Studies in the Linguistic Sciences* 10(1):71–84
- Jelinek, Eloise and Carnie, Andrew 2003. 'Argument hierarchies and the mapping principle.' In Carnie, Andrew, Harley, Heidi, and Willie, Mary Ann (eds.) 'Formal Approaches to Function in Grammar: In Honor of Eloise Jelinek,' Philadelphia: John Benjamins, pp. 265–296
- Johnson, David E. and Postal, Paul M. 1980. *Arc Pair Grammar*. Princeton: Princeton University Press
- Kadmon, Nirit 2001. *Formal Pragmatics*. Malden: Blackwell Publishers
- Kamp, Hans and Reyle, Uwe 1993. *From Discourse to Logic: An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers
- Kaplan, Ronald M. 1987. 'Three seductions of computational psycholinguistics.' In Whitelock, Peter, Wood, Mary McGee, Somers, Harold L., Johnson, Rod, and Bennett, Paul (eds.) 'Linguistic Theory and Computer Applications,' London: Academic Press, pp. 149–188. Also published as CCL/UMIST Report No. 86.2: Alvey/ICL Workshop on Linguistic Theory and Computer Applications: Transcripts of Presentations and Discussions. Center for Computational Linguistics, University of Manchester. Reprinted in Dalrymple et al. (1995:337–367)
- Kaplan, Ronald M. 1988. 'Correspondences and their inverses.' Presented at the Titisee Workshop on Unification Formalisms: Syntax, Semantics and Implementation, Titisee, Germany
- Kaplan, Ronald M. and Bresnan, Joan 1982. 'Lexical-Functional Grammar: A formal system for grammatical representation.' In Bresnan, Joan (ed.) 'The Mental Representation of Grammatical Relations,' Cambridge, MA: The MIT Press, pp. 173–281. Reprinted in Dalrymple et al. (1995:29–130)

- Kaplan, Ronald M. and Zaenen, Annie 1989. 'Long-distance dependencies, constituent structure, and functional uncertainty.' In Baltin, Mark and Kroch, Anthony (eds.) 'Alternative Conceptions of Phrase Structure,' Chicago University Press, pp. 17–42. Reprinted in Dalrymple et al. (1995:137–165)
- Karimi, Simin 1990. 'Obliqueness, specificity, and discourse functions: Râ in Persian.' *Linguistic Analysis* 20:139–191
- Keenan, Edward L. 1976. 'Towards a universal definition of subject.' In Li, C. N. (ed.) 'Subject and Topic,' New York: Academic Press, pp. 303–333
- Keenan, Edward L. and Comrie, Bernard 1977. 'Noun phrase accessibility and universal grammar.' *Linguistic Inquiry* 8(1):63–99
- Kellogg, Samuel Henry 1955 [1893]. *A grammar of the Hindi language*. London: Routledge and Kegan Paul
- Kepping, Ksenia B. 1979. 'Elements of ergativity and nominativity in Tangut.' In Planck, Frans (ed.) 'Ergativity,' London: Academic Press, pp. 263–277
- Kibrik, Aleksandr and Seleznev, A. 1980. 'Sintaksis i morfologija glagol' nogo soglasovanija v tabasaranskom jazyke [The syntax and morphology of verbal agreement in Tabassaran].' In Zvegincev, Vladimir (ed.) 'Tabasaranskije etjudy. Materialy dagestanskoj ekspedicii 1979 g.', Moscow: MGU, pp. 17–33
- Kifle, Nazareth Amlesom 2007. 'Differential object marking and topicality in Tigrinya.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2007 Conference,'
- Kim, Alan 1988. 'Preverbal focus position in type XIII languages.' In Hammond, Michael, Moravcsik, Edith, and Wirth, Jessica (eds.) 'Studies in syntactic typology,' Amsterdam: John Benjamins, pp. 148–171
- King, Tracy Holloway 1995. *Configuring Topic and Focus in Russian*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Revised and corrected version of 1993 Stanford University dissertation.
- King, Tracy Holloway 1997. 'Focus domains and information structure.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG97 Conference,' URL <http://csli-publications.stanford.edu/LFG/2/lfg97.html>

- King, Tracy Holloway and Zaenen, Annie 2004. 'F-structures, information structure, and discourse structure.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2004 Conference,' URL <http://csli-publications.stanford.edu/LFG/9/lfg04.html>
- Kittilä, Seppo 2006a. 'Object-, animacy- and role-based strategies: A typology of object marking.' *Studies in Language* 30:1–32
- Kittilä, Seppo 2006b. 'The woman showed the baby to her sister: On resolving humanness-driven ambiguity in ditransitives.' In Kulikov, Leonid, Malchukov, Andrej, and de Swart, Peter (eds.) 'Case, Valency and Transitivity,' Amsterdam: John Benjamins, pp. 291–308
- Kornfilt, Jaklin 2009. 'DOM and two types of DSM in Turkish.' In de Hoop, Helen and de Swart, Peter (eds.) 'Differential Subject Marking,' Dordrecht: Springer, pp. 79–111
- Krifka, Manfred 1992. 'A framework for focus-sensitive quantification.' In Barker, Chris and Dowty, David (eds.) 'The Proceedings from the Second Conference on Semantics and Linguistic Theory: SALT II,' Department of Linguistics, Ohio State University, pp. 215–236. Ohio State University Working Papers in Linguistics, Number 40
- Krifka, Manfred 2006. 'Association with focus phrases.' In Molnar, Valerie and Winkler, Susanne (eds.) 'The Architecture of Focus,' Berlin: Mouton de Gruyter
- Kroeger, Paul 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Revised and corrected version of 1991 Stanford University dissertation.
- Kulonen, Ulla-Maja 1989. *The passive in Ob-Ugrian*. Helsinki: Finno-Ugrian Society
- Kuznecova, Ariadna, Xelimskij, Evgenij, and Gruskina, Elena 1982. *Očerki po sel'kupskomu jazyku [Studies in Selkup]*. Moscow: MGU
- Kwon, Song-Nim and Zribi-Hertz, Anne 2008. 'Differential function marking, case and information structure: Evidence from Korean.' *Language* 84(2):218–257
- Lam, Olivia S.-C. 2008. *Object Functions and the Syntax of Double Object Constructions in Lexical Functional Grammar*. D.Phil. dissertation, Oxford University

- Lambrecht, Knud 1981. *Topic, antitopic and verb agreement in non-standard French*. Amsterdam: John Benjamins
- Lambrecht, Knud 1987. 'On the status of SVO sentences in French discourse.' In Tomlin, Russell S. (ed.) 'Coherence and Grounding in Discourse,' Amsterdam and Philadelphia: John Benjamins, pp. 217–261
- Lambrecht, Knud 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge, UK: Cambridge University Press
- Lambrecht, Knud 2000. 'When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages.' *Studies in Language* 24:611–682
- Lambrecht, Knud and Polinsky, Maria 1997. 'Typological variation in sentence focus constructions.' In 'Papers from the Thirty-Third Regional Meeting of the Chicago Linguistic Society: Panels on Linguistic Ideologies in Contact,' University of Chicago: Chicago Linguistic Society, pp. 189–206
- Lazard, Gilbert 1982. 'Le morphème *râ* en persan et les relations actanciennes.' *Bulletin de la société de linguistique de Paris* 73:177–208
- Lazard, Gilbert 1984. 'Actance variations and categories of the object.' In Plank, Frans (ed.) 'Objects: Towards a theory of grammatical relations,' London: Academic Press, pp. 269–292
- Lazard, Gilbert 1992. *A grammar of contemporary Persian*. Costa Mesa, CA: Mazda Publishers
- Legendre, Géraldine, Vikner, Sten, and Grimshaw, Jane (eds.) 2001. *Optimality-Theoretic Syntax*. The MIT Press
- Leonetti, Manuel 2003. 'Specificity and object marking: the case of Spanish *a*.' In von Stechow, Klaus and Kaiser, Georg A. (eds.) 'Proceedings of the Workshop *Semantic and syntactic aspects of specificity in Romance languages*,' Konstanz: University of Konstanz, pp. 67–101
- Levin, Lori S. 1986. *Operations on Lexical Forms: Unaccusative Rules in Germanic Languages*. Ph.D. thesis, MIT
- Liao, Gwen W.-T. 2010. *An LFG Account of Empty Pronouns in Mandarin Chinese*. D.Phil. dissertation, Oxford University
- Lidz, Jeffrey 2006. 'The grammar of accusative case in Kannada.' *Language* 82(1):10–32

- Lødrup, Helge 1999. 'Linking and optimality in the Norwegian presentational focus construction.' *Nordic Journal of Linguistics* 22(2):205–229
- MacKay, Carolyn J. and Trechsel, Frank R. 2008. 'Symmetrical objects in Misantla Totonac.' *International Journal of American Linguistics* 74(2):227–255
- Mahootian, Shahrzad 1997. *Persian*. London and New York: Routledge
- Malchukov, Andrej 2006. 'Transitivity parameters and transitivity alternations: Constraining co-variation.' In Kulikov, L. I., Malchukov, Andrej, and de Swart, Peter (eds.) 'Case, valency, and transitivity: A cross-linguistic perspective,' Amsterdam: John Benjamins, pp. 330–357
- Malchukov, Andrej, Haspelmath, Martin, and Comrie, Bernard 2007. 'Ditransitive constructions: A typological overview.' URL <http://email.eva.mpg.de/~haspelmt/DitransitiveOverview.pdf>. Draft for the Conference on Ditransitive Constructions, November 2007
- Malchukov, Andrej L. 2008. 'Animacy and asymmetries in differential case marking.' *Lingua* 118:203–221
- Maling, Joan and Zaenen, Annie (eds.) 1990. *Syntax and Semantics: Modern Icelandic Syntax*, volume 24. San Diego: Academic Press
- Marcantonio, Angela 1985. 'On the definite vs. indefinite conjugation in Hungarian: A typological and diachronic analysis.' *Acta Linguistica Academiae Scientiarum Hungaricae* 35:267–298
- Masica, Colin 1982. 'Identified object marking in Hindi and other languages.' In Koul, O. N. (ed.) 'Topics in Hindi linguistics,' New Delhi: Bahri Publications, pp. 16–50
- Masica, Colin 1991. *The Indo-Aryan languages*. Cambridge, UK: Cambridge University Press
- Maslova, Elena 2003a. *A Grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter
- Maslova, Elena 2003b. 'Information focus in relational clause structure.' In Tsunoda, Tasaku and Kageyama, Taro (eds.) 'Voice and Grammatical Relations: Festschrift for Masayoshi Shibatani,' Amsterdam: John Benjamins, Typological Studies in Language, pp. 175–194



- Matsumoto, Yo 1996. *Complex Predicates in Japanese: A Syntactic and Semantic Study of the Notion 'Word'*. Ph.D. thesis, Stanford University. *Studies in Japanese Linguistics* series, CSLI Publications/Kuroiso Publishers, Stanford/Tokyo. Revised and corrected version of 1992 Stanford University dissertation, *On the wordhood of complex predicates in Japanese*.
- Mayer, Elisabeth 2008. 'Clitics on the move: From dependent marking to split marking.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2008 Conference,' URL <http://csli-publications.stanford.edu/LFG/13/lfg08.html>
- McCawley, James D. 1968. 'Concerning the base component of a transformational grammar.' *Foundations of Language* 4:243–269. Reprinted in McCawley (1973)
- McCawley, James D. 1973. *Grammar and Meaning: Papers on Syntactic and Semantic Topics*. Tokyo: Taishukan
- McGregor, R. S. 1972. *Outline of Hindi Grammar*. Oxford/Delhi: Oxford University Press
- Meinunger, André 1998. 'Topicality and agreement.' In Darnell, Michael, Moravcsik, Edith A., Noonan, Michael, Newmeyer, Frederick J., and Wheatley, Kathleen M. (eds.) 'Functionalism and formalism in linguistics,' Amsterdam: John Benjamins, volume 2, pp. 203–219
- Mithun, Marianne 1991. 'Active/agentive case marking and its motivation.' *Language* 67:510–546
- Mohanan, K. P. 1982. 'Grammatical relations and clause structure in Malayalam.' In Bresnan, Joan (ed.) 'The Mental Representation of Grammatical Relations,' Cambridge, MA: The MIT Press, pp. 504–589
- Mohanan, Tara 1994. *Arguments in Hindi*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Reprinted version of 1990 Stanford University dissertation.
- Morimoto, Yukiko 2002. 'Prominence mismatches and differential object marking in Bantu.' In Butt, Miriam and King, Tracy Holloway (eds.) 'On-line Proceedings of the LFG2002 Conference,' URL <http://csli-publications.stanford.edu/LFG/7/lfg2.html>
- Morimoto, Yukiko 2009. 'From topic to subject marking: Implications for a typology of subject marking.' In de Hoop, Helen and de Swart, Peter (eds.) 'Differential Subject Marking,' Dordrecht: Springer, pp. 199–221

- Mycock, Louise 2006. *The Typology of Constituent Questions: A Lexical Functional Grammar analysis of 'wh'-questions*. Ph.D. thesis, University of Manchester
- Mycock, Louise 2009. "‘What do you do?’: Variation in interrogative predicates." Paper presented at the workshop "Blurring Component Boundaries: levels of analysis or growth of information?", LFG09, Cambridge, July 2009
- Næss, Åshild 2004. 'What markedness marks: The markedness problem with direct objects.' *Lingua* 114:1186–1212
- Næss, Åshild 2006. 'Case semantics and the agent-patient opposition.' In Kulikov, Leonid, Malchukov, Andrej, and de Swart, Peter (eds.) 'Case, valency and transitivity,' Amsterdam: John Benjamins, pp. 309–327
- Næss, Åshild 2007. *Prototypical Transitivity*. Amsterdam: John Benjamins
- Neeleman, Ad and van de Koot, Hans 2008. 'Dutch scrambling and the nature of discourse templates.' *Journal of Comparative Germanic Linguistics* 11(2):137–189
- Nichols, Johanna 1986. 'Head-marking and dependent-marking grammar.' *Language* 62(1):56–119
- Nikolaeva, Irina 1999. 'Object agreement, grammatical relations, and information structure.' *Studies in Language* 23:331–376
- Nikolaeva, Irina 2001. 'Secondary topic as a relation in information structure.' *Linguistics* 39:1–49
- Nordlinger, Rachel 1998. *Constructive Case: Evidence from Australian Languages*. Dissertations in Linguistics. Stanford, CA: CSLI Publications. Revised version of 1997 Stanford University dissertation.
- O'Connor, Robert 2006. *Information Structure in Lexical Functional Grammar: The Discourse-Prosody Correspondence*. Ph.D. thesis, University of Manchester
- Paoli, Sandra 2009. 'Contrastiveness and new information: A new view on focus.' ms, University of Oxford
- Pápay, József 1906-1908. 'Északi-osztják nyelvtanulmányok [Studies on Northern Ostyak].' *Nyelvtudományi Közlemények* 36–38:345–398, 52–79, 164–195, 258–275, 111–150, 313–329

- Paul, Ludwig 2003. 'Early Judaeo-Persian in a historical perspective: The case of the prepositions *be*, *u*, *pa(d)*, and the suffix *rā*.' In Paul, Ludwig (ed.) 'Persian Origins - Early Judaeo-Persian and the emergence of New Persian,' Wiesbaden: Harrassowitz Verlag, pp. 177–194
- Payne, Doris L. and Barshi, Immanuel 1999. 'External possession: What, how, and why.' In Payne, Doris L. and Barshi, Immanuel (eds.) 'External possession,' Amsterdam: John Benjamins, pp. 3–29
- Peterson, David A. 2007. *Applicative Constructions*. Oxford: Oxford University Press
- Plank, Frans (ed.) 1995. *Double Case: Agreement by Suffixaufnahme*. Oxford: Oxford University Press
- Platzack, Christer 1983. 'Existential sentences in English, Swedish, German and Icelandic.' In Karlsson, Fred (ed.) 'Papers from the Seventh Scandinavian Conference of Linguistics,' University of Helsinki, Department of General Linguistics, pp. 80–100
- Polinsky, Maria 1995. 'Double objects in causatives: Towards a study of coding conflict.' *Studies in Language* 19:129–221
- Polinsky, Maria 1998. 'A non-syntactic account of some asymmetries in the double object construction.' In Koenig, Jean-Pierre (ed.) 'Conceptual structure and language: Bridging the gap,' Stanford, CA: CSLI Publications, pp. 403–423
- Polinsky, Maria and Potsdam, Eric 2001. 'Long-distance agreement and topic in Tsez.' *Natural Language and Linguistic Theory* 19(3):583–646
- Portner, Paul and Yabushita, Katsuhiko 1998. 'The semantics and pragmatics of topic phrases.' *Linguistics and Philosophy* 21(2):117–157
- Portner, Paul and Yabushita, Katsuhiko 2001. 'Specific indefinites and the information structure theory of topics.' *Journal of Semantics* 18:271–297
- Prince, Ellen 1981. 'Topicalization, focus-movement, and Yiddish-movement: a pragmatic differentiation.' In 'Proceedings of the Seventh Annual Meeting of the Berkeley Linguistics Society,' Berkeley Linguistics Society
- Ramchand, Gillian 1997. *Aspect and Predication: The Semantics of Argument Structure*. Oxford University Press
- Raz, Shlomo 1983. *Tigre grammar and texts*. Malibu: Undena Publications

- Rédei, Károly 1965. *Northern Ostyak chrestomathy*. The Hague, Bloomington: Indiana University Press
- Reinhart, Tanya 1982. 'Pragmatics and linguistics: An analysis of sentence topics.' *Philosophica* 27:53–94
- Ritter, Elizabeth and Rosen, Sara Thomas 2001. 'The interpretive value of object splits.' *Language Sciences* 23:425–451
- Rizzi, Luigi 1997. 'The fine structure of the left periphery.' In Haegeman, Liliane (ed.) 'Elements of Grammar,' Dordrecht: Kluwer Academic Publishers, pp. 281–337
- Roberts, Craig 1996. 'Information structure in discourse: Towards an integrated formal theory of pragmatics.' In Kathol, Andreas and Yoon, Jae-Hak (eds.) 'Ohio State University Working Papers in Linguistics,' Ohio State University, volume 49, pp. 91–136
- Rooth, Mats 1992. 'A theory of focus interpretation.' *Natural Language Semantics* 1:75–116
- Rooth, Mats E. 1985. *Association with Focus*. Ph.D. thesis, University of Massachusetts at Amherst
- Rude, Noel 1982. 'Promotion and topicality of Nez Perce objects.' In Macaulay, Monica, Gensler, Orin D., Brugman, Claudia, Civkulis, Inese, Dahlstrom, Amy, Krile, Katherine, and Sturm, Rob (eds.) 'Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society,' University of California, Berkeley: Berkeley Linguistics Society, pp. 463–483
- Rude, Noel 1986. 'Topicality, transitivity, and the direct object in Nez Perce.' *International Journal of American Linguistics* 52(2):124–153
- Saeed, John I. 1984. *The syntax of topic and focus in Somali*. Hamburg: Helmut Buske
- Saeed, John I. 1987. *Somali Reference Grammar*. Wheaton, MD: Dunwoody
- Sag, Ivan A., Wasow, Thomas, and Bender, Emily 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications
- Sasse, Hans-Jurgen 1984. 'The pragmatics of noun incorporation in Eastern Cushitic.' In Plank, Frans (ed.) 'Objects: Towards a theory of grammatical relations,' London: Academic Press, pp. 243–268
- Schulz, B. 2005. 'Tune in – drop out: Harmonic alignment as a prerequisite for German topic drop.' Cited by Erteschik-Shir (2007:24)

- Schwarzschild, Roger 1997. 'Why some foci must associate.' MS, Rutgers University
- Seidl, Amanda and Dimitriadis, Alexis 1997. 'The discourse function of object marking in Swahili.' In 'Papers from the Thirty-Third Regional Meeting of the Chicago Linguistic Society: Panels on Linguistic Ideologies in Contact,' University of Chicago: Chicago Linguistic Society, pp. 373–387
- Siewierska, Anna 2004. *Person*. Cambridge, UK: Cambridge University Press
- Silverstein, Michael 1976. 'Hierarchy of features and ergativity.' In Dixon, R.M.W. (ed.) 'Grammatical Categories in Australian Languages,' Canberra: Australian Institute for Aboriginal Studies, pp. 112–171
- Simpson, Jane 1991. *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Dordrecht: Kluwer Academic Publishers
- Singh, Mona 1994. 'Thematic roles, word order, and definiteness.' In Butt, Miriam, King, Tracy Holloway, and Ramchand, Gillian (eds.) 'Theoretical perspectives on word order in South Asian languages,' Stanford, CA: CSLI Publications, pp. 217–235
- Skribnik, Elena 2001. 'Pragmatic structuring in Northern Mansi.' In Seilenthal, Tõnu (ed.) 'Congressus nonus internationalis fenno-ugristarum. 7.-13.8.2000 Tartu. Pars IV. Dissertationes sectionum: Linguistica III,' Tartu: Tartu University, pp. 222–239
- Sportiche, Dominique 1995. 'Clitic constructions.' In Rooryck, Johan and Zaring, Laurie (eds.) 'Phrase Structure and the Lexicon,' Dordrecht: Kluwer Academic Publishers, pp. 213–276
- Steedman, Mark J. 2001. *The Syntactic Process*. Cambridge, MA: The MIT Press
- Stowell, Tim, Wehrli, Eric, and Anderson, Stephen R. (eds.) 1992. *Syntax and Semantics: Syntax and the Lexicon*, volume 26. San Diego: Academic Press
- Strawson, Paul F. 1964. 'Identifying reference and truth-values.' In Steinberg, Danny D. and Jakobovitz, Leon A. (eds.) 'Semantics: An interdisciplinary reader,' Cambridge, UK: Cambridge University Press, pp. 86–99
- Strawson, Paul F. 1974. *Subject and predicate in logic and grammar*. London: Ashgate

- Stump, Gregory T. and Yadav, Ramawatar 1988. 'Maithili verb agreement and the Control Agreement Principle.' In 'Papers from the 24th Regional Meeting of the Chicago Linguistic Society, Part 2: Parasession on Agreement in Grammatical Theory,' volume 24, pp. 304–321
- Szabolcsi, Anna 1981. 'The semantics of topic-focus articulation.' In Groenendijk, Jeroen, Janssen, Theo, and Stokhof, Martin (eds.) 'Formal Methods in the Study of Language,' Amsterdam: Mathematical Centre, pp. 513–540
- Tereshchenko, Natal'ja 1979. *Nganasanskij jazyk [The Nganasan language]*. Leningrad: Nauka
- Toldova, Svetlana and Serdobolskaya, Natalia 2008. 'Differential object marking in Uralic languages: a multiparameter analysis.' Paper presented at *Cognitive and Functional Perspectives on Dynamic Tendencies in Languages*, University of Tartu, Estonia, May 29–June 1, 2008
- Torrego, Esther 1998. *The dependencies of objects*. Cambridge, MA: The MIT Press.
- Vallduví, Enric 1992. *The Informational Component*. New York: Garland
- Vallduví, Enric and Engdahl, Elisabet 1996. 'The linguistic realization of information packaging.' *Linguistics* 34:459–519
- van Geenhoven, Veerle 1998. *Semantic incorporation and indefinite descriptions: Semantic and syntactic aspects of noun incorporation in West Greenlandic*. Stanford, CA: CSLI Publications
- Van Valin, Robert D. 2003. *Exploring the Syntax-Semantics Interface*. Cambridge, UK: Cambridge University Press
- Vincent, Nigel 1997. 'LFG as a model of syntactic change.' In Butt, Miriam and King, Tracy Holloway (eds.) 'Time over Matter: Diachronic Perspectives on Morphosyntax,' Stanford, CA: CSLI Publications, pp. 1–42
- von Heusinger, Klaus and Kornfilt, Jaklin 2005. 'The case of the direct object in Turkish: Semantics, syntax and morphology.' *Turkic Languages* 3:3–44
- von Stechow, Arnim 1982. 'Structured propositions.' Report of the Konstanz SFB. URL <http://www2.sfs.uni-tuebingen.de/~arnim10/Aufsaetze/Structured%20Prop%201.pdf>
- Windfuhr, Gernot 1979. *Persian grammar: History and state of its study*. Trends in linguistics. The Hague: Mouton de Gruyter

- Woolford, Ellen 1999. 'Animacy hierarchy effects on object agreement.' In Kotey, Paul (ed.) 'New dimensions in African linguistics and languages (Trends in African Linguistics 3),' Trenton: Africa World Press, pp. 203–216
- Woolford, Ellen 2000. 'Object agreement in Palauan: Specificity, humanness, economy and optimality.' In Paul, Ileana, Phillips, Vivianne, and Travis, Lisa (eds.) 'Formal issues in Austronesian linguistics,' Dordrecht: Kluwer Academic Publishers, pp. 215–245
- Woolford, Ellen 2001. 'Conditions on object agreement in Ruwund (Bantu).' In Benedicto, E. (ed.) 'The UMass volume on indigenous languages (University of Massachusetts Occasional Papers in Linguistics 20),' Amherst, MA: GLSA. URL <http://people.umass.edu/ellenw/Woolford%20Conditions%20on%20object%20Agreement.pdf>
- Xelimskij, Evgenij A. 1982. *Drevnejshie vengersko-samodijskie jazykovye paralleli [Ancient linguistic parallels between Hungarian and Samoyed]*. Moscow: Nauka
- Yamamoto, Mutsumi 1999. *Animacy and reference*. Amsterdam: John Benjamins
- Yang, Ning and van Bergen, Geertje 2007. 'Scrambled objects and case marking in Mandarin Chinese.' *Lingua* 117:1617–1635
- Zaenen, Annie, Maling, Joan, and Thráinsson, H. 1985. 'Case and grammatical functions: The Icelandic passive.' *Natural Language and Linguistic Theory* 3(4):441–483. Reprinted in Maling and Zaenen (1990:95–164)
- Zwart, Jan-Wouter 2009. 'Uncharted territory? Towards a non-cartographic account of Germanic syntax.' In Alexiadou, Artemis, Hankamer, Jorge, McFadden, Thomas, Nuger, Justing, and Schäfer, Florian (eds.) 'Advances in Comparative Germanic Syntax,' Amsterdam: John Benjamins