Subjects and Universal Grammar

An Explanatory Theory

Yehuda N. Falk

This page intentionally left blank

SUBJECTS AND UNIVERSAL GRAMMAR

The "subject" of a sentence is a concept that presents great challenges to linguists. Most languages have something which looks like a subject, but subjects differ across languages in their nature and properties, making them an interesting phenomenon for those seeking linguistic universals. This pioneering volume takes a new approach to subjects, addressing their nature from a simultaneously formal and typological perspective. Dividing the subject into two distinct grammatical functions, it shows how the nature of these functions explains their respective properties, and argues that the split in properties shown in "ergative" languages (whereby the subject of intransitive verbs is marked as an object) results from the functions being assigned to different elements of the clause. Drawing on data from a typologically wide variety of languages, and examining a range of constructions, this book explains why, even in the case of very different languages, certain core properties can be found.

YEHUDA N. FALK is Senior Lecturer in the Department of English, The Hebrew University of Jerusalem. He has also been a Visiting Scholar at Stanford University (1999–2000). He has previously published *Lexical-Functional Grammar: An Introduction to Parallel, Constraint-based Syntax* (2001), and has contributed to a variety of journals including *Language* and the *Journal of Linguistics*.

In this series

Earlier issues not listed are also available

Subjects and Universal Grammar

An Explanatory Theory

SUBJECTS AND UNIVERSAL GRAMMAR

AN EXPLANATORY THEORY

YEHUDA N. FALK

The Hebrew University of Jerusalem

To the people responsible for my linguistic career:
   my late grandmother
   Barbara Klima ז״ל
who first introduced me to a language other than English

   and my parents
   Paul and Eva Falk יבל״א
who encouraged my mishegas with languages.

(Who would have thought that those Berlitz records would lead to this?)

Contents

Preface

According to the biblical book of Kohelet (Ecclesiastes), "the making of many books is without end." I don't know about "many books," but the making of this book has sometimes appeared to be without end. It began some forgotten day in the late 1980s when the idle thought crossed my mind: "How might one redesign GB Case theory to account for ergative languages?" A very early exploration of the issues in this book, in the guise of GB Case theory, was published in *Linguistics* in 1991, under the title "Case: Abstract and Morphological." I also presented several papers on Case, ergativity, and such at conferences of the Israel Association for Theoretical Linguistics in the 1990s. But in the course of trying to understand ergative languages I began to realize that the GB framework was missing something. What this "something" was started to become clearer to me when I started considering Philippine-type languages, because it was obvious to me that direct reference to grammatical functions was necessary to account for the "voice" morphology.

This realization led me back to LFG, the theoretical framework in which I had begun my linguistic career. I began reframing the work that I had been doing in terms of LFG. A presentation at the 1999 conference of the Austronesian Formal Linguistics Association received encouraging responses. In the fall semester of the 1999–2000 academic year, I was fortunate to be able to spend a sabbatical as a Visiting Scholar at Stanford University, hosted by Joan Bresnan. I spent incredible amounts of time in the Green Library with my laptop, taking notes from books not available in Jerusalem. Based on my reading, and with enthusiastic encouragement from Joan, I started focusing my attention on issues of subjecthood, and started to take seriously languages I hadn't considered before and constructions that I didn't really understand earlier. It was also as a result of Joan's encouragement that I began to think of writing a book. The core of this book was presented at the LFG 2000 conference, and I have presented this material in departmental colloquia at the Hebrew University of Jerusalem and Tel Aviv University. However, this book took a back seat to another project that grew out of my sabbatical, my LFG textbook *Lexical-Functional Grammar: An*

Introduction to Parallel Constraint-Based Syntax (2001). Finally I was able to complete the manuscript, only to be faced with two major rewrites as a result of comments by readers for Cambridge University Press.

There are many people who have had a hand in helping me complete this book. In the first place, this book would not exist without the fieldworkers who have collected the data on which this book is based. While I know very few of them personally, I am forever indebted to these hardy souls. They have forever enriched the database on which linguistics works and, if generative linguistics is to be the search for the nature of Universal Grammar, it is only through their continued efforts that the field will be able to progress.

Joan Bresnan, as my sponsor at Stanford, has provided immeasurable input into this study and much invaluable moral support, as well as being my role model as a descriptive/theoretical linguist. Ron Kaplan, the keeper of the LFG formalism, helped me out on a couple of occasions when I couldn't find the right way to express something. My Hebrew University colleague Yael Ziv has helped me realize the importance of pragmatics in language, and given me a new appreciation for the insights (if not the formulations) of functionalists. Ray Jackendoff, from whom I first learned transformational syntax back in 1976, encouraged me to abandon the transformational model; the influence of his views on language should be apparent to all. Other people who have commented on portions of the material here and/or helped me with data include Alex Alsina, I Wayan Arka, Aaron Broadwell, Elizabeth Coppock, Mary Dalrymple, Edit Doron, Mike Dukes (my officemate at Stanford), Fred Landman, Paul Kroeger, Chris Manning, Irit Meir, Anita Mittwoch, Asya Pereltsvaig, Ivan Sag, Jane Simpson, the late Joe Taglicht, Lisa Travis, and Annie Zaenen. Participants in the conferences and colloquia at which I have presented this material have been very helpful. I am also grateful to the students who have taken graduate seminars in which this material has been covered in one form or another, in 1996, 1998, 2000, 2001, and 2003. In teaching the material to them I was better able to formulate the half-baked ideas that were swirling around in my mind. Andrew Winnard at Cambridge University Press has been very helpful and encouraging, and the Press's anonymous readers forced me to go over the material again and again, first fleshing out the points I have tried to make, then strengthening the argumentation. Thanks to their comments, the book is much better than the first manuscript that I submitted. For making my Stanford sabbatical more enjoyable, I would like to thank the Palo Alto Jewish community, especially Rabbi Shelly Lewis and the rest of the folks at Congregation Kol Emeth.

My wife Brandel, a longtime La Leche League leader, has, as always, been an inspiration with her dedication. My sons, all now either in or approaching

adulthood, have enriched my life in various ways which have helped me complete this project: Eli with his interest in academic endeavor; Yoni with his fierceness of conviction; Mati with his unbridled enthusiasm; and Gabi with his still-open-minded childlike innocence. And my baby daughter Pnina has helped me rediscover what an incredible journey of exploration life is (and how fascinating language is).

My maternal grandmother, Barbara Klima, passed away while this book was under review. A survivor of the Holocaust, she made a new life for herself and her daughter (my mother), and lived to age 98, seeing seven great-grandchildren. Her strength of spirit was inspirational. I miss her terribly.

Notes on the text

Dyirbal and Yidin[y] examples are presented using the practical orthography currently employed by Australianists (as in Dixon 1994). The examples from Dixon's grammars (1972 for Dyirbal and 1977 for Yidin[y]) have been updated accordingly: ɲ has been changed to n[y], ḍ to j, ṛ to r, and r to rr.

Except for section 1.2.2 on Case marking, absolutive and nominative case are only glossed when there are overt markers.

I follow the typographical convention of capitalizing the word "Case". This notation was introduced in early Government/Binding theory as a device for disambiguating the word "case", a word which happens to have a wide-ranging set of meanings: "I will follow the practice of capitalizing 'Case' when it is used in the technical sense, to avoid confusion with informal use, as in 'the unmarked case,' etc." (Chomsky 1980: 13 fn. 18). The distinction is a useful one; in fact, taking Chomsky's own example, one wants to distinguish between "unmarked case" (i.e. unmarked situation) and "unmarked Case" (unmarked morphological form of a noun). It is in this spirit that the capitalization is being used here. This notation has, over the years, acquired an unfortunate sense of distinguishing some abstract, theoretical notion of Case from ordinary morphological Case. In the present study, Case refers to morphological marking.

Abbreviations used in glosses

numbers	(in examples from Bantu languages) noun class
ABL	ablative case
ABS	absolutive case or agreement
ACC	accusative case
ACT	actor "voice" (nominative = A argument) in Philippine-type languages
ADJ	adjective
ADNOM	adnominal
AGT	agent agreement
ALL	allative case
APASS	antipassive
APPL	applicative
ASP	aspectual marker
AUX	auxiliary
BEN	benefactive "voice" (nominative = benefactive) in Philippine-type languages
CAUS	causative
CLASS	classifier
CNTMP	contemplated tense
COMP	complementizer
COMPL	completive
DAT	dative case
DECL	declarative
DEF	definite
DEICT	deictic
DIFF	different subject (in switch-reference systems)
DIR	directional
DIRS	directional suffix
DIST	distal realis

DO	direct-object "voice" (nominative = P argument) in Philippine-type languages
DU	dual
ERG	ergative case or agreement
EXCL	exclusive
F	feminine
FSG	feminine singular
FOC	focus
FUT	future tense
GEN	genitive case
GER	gerund
IMM	immediate
IMP	imperative
IMPERF	imperfect(ive)
IMPLIC	implicated clause
INCH	inchoative
INCL	inclusive
IND	indicative
INF	infinitive
INS	instrumental "voice" (nominative = instrument) in Philippine-type languages
INSTR	instrumental case
INTR	intransitive
IO	indirect-object "voice" (nominative = indirect object, locative, or directional) in Philippine-type languages
IRR	irrealis
LNK	linker
LOC	locative case
M	masculine
MSG	masculine singular
NEG	negative
NFUT	non-future
NMNL	nominalizer
NOM	nominative case
NONVOL	non-volitive mood
NPST	non-past
OBJ	object agreement marker or case
OBL	oblique
OCONTR	object "control" in Walpiri

PART	participle
PASS	passive
PAT	patient agreement
PERF	perfect(ive)
PERS	noun referring to a person
PL	plural
POL	polite
POSS	possessive
PRES	present tense
PRON	pronoun
PROP	proper noun
PSPRT	passive participle
PST	past tense
Q	question
REAL	realis
REC	recent past tense
RECIP	reciprocal
REFL	reflexive
REL	relative
SAME	same subject (in switch-reference systems)
SBJCT	subjunctive
SCONTR	subject "control" (same subject) in Warlpiri
SG	singular
STAT	stative
SUBJ	subject agreement marker
SUFF	suffix
TNS	tense
TOP	topic
TR	transitive
TRANSL	translative case
VWL	vowel (thematic or similar phonological augment)

1   *On subjects and explanation*

1.1   Overview

Explaining subjects and their properties is an important challenge in contemporary linguistics. For formalist approaches to linguistics, the clustering of properties that subjects display necessitates some special representational properties unique to subjects. Without such representational uniqueness, the properties of subjects that set them apart from other elements of the clause are mysterious. However, this only pushes the need for explanation back one level: such special representation itself calls out for explanation. For functionalist approaches, similar issues are raised, as it is not clear what the functional properties of subjects are that set them apart. From a typological perspective, the mystery of subjects is even deeper, as different language types appear to deploy subject properties in different (but systematic) ways. As a result of the discoveries of ergative languages, Philippine-type languages, active languages, and the like, interesting questions have been raised about the properties of subjects, the representation of subjects, and even the cross-linguistic validity of "subject" as an element of linguistic description.

The concept of "subject" is one with a long history in linguistics. As with most other such concepts, contemporary linguistics did not invent the subject. Instead, it has taken a traditional concept and attempted to provide it with theoretical content. Problems have arisen because the concept "subject" originates in traditional studies of classical Indo-European languages such as Greek and Latin, languages which are closely related genetically, areally, and typologically. Investing "subject" with theoretical content thus usually depends on either focusing on languages which are typologically similar to classical Indo-European languages or attempting to extend an Indo-European notion to languages which have very different typological properties. As a result, different researchers take varying positions on which languages are examined, and in some languages which element (if any) is to be identified as the subject. Much of the literature on such topics as ergativity and active languages focuses on

debates such as these. These issues need to be clarified if a true understanding of subjects and their properties is to be achieved.

All contemporary approaches to linguistics – formalist, functionalist, typological, etc. – appropriately take the goal of linguistics to be the explanation of linguistic phenomena. As such, they depart from merely being satisfied with describing linguistic facts, although proper description is, of course, a prerequisite for explanation. In the realm of subjecthood, this means that simply stipulating the properties of subjects is not sufficient: the properties should follow from a proper characterization of the nature of subjects. Since explanation is possible only in the context of a theory of the linguistic domain in question, the attempts that have been made at explaining subjects have been as varied as schools of linguistics, and have mirrored the drawbacks of the theoretical assumptions made by the researchers. Formal accounts tend to be characterized by a disregard for functional factors and often by inadequate cross-linguistic coverage. Functionalist and typological accounts are typically based on superficial surveys of languages and disregard the nature of the formal devices involved in syntax.

It is the thesis of this study that a truly explanatory theory of subjects has yet to be constructed, and its goal is the proposal of such a theory. A theory of subjects must be formally grounded, functionally aware, and achieve sufficiently broad typological coverage, including all of the types of languages which are potentially problematic. Unlike previous accounts, the theory of subjects to be proposed here meets all of these criteria. Naturally, it draws on insights of earlier approaches, but it synthesizes them in a way which results in true explanation of the properties of subjects as they are revealed in cross-linguistic study.

In this first chapter, we will enumerate the properties generally thought to be subject properties. We will also discuss typological issues related to subjecthood. Finally, we will discuss different types of approaches to subjects.

1.2   Subject properties

1.2.1   *First approximation*

As mentioned above, subjects display an array of properties which must be accounted for by a theory of subjecthood. Properties of subjects have been enumerated in studies like Keenan (1976) and Andrews (1985). We will review them here briefly, primarily using examples from English. However, before we discuss the properties of subjects, it is necessary to take heed of the following observation by Andrews (1985: 104):

> At the outset we must note that there are no properties which in all languages
> are always exhibited by subjects and only exhibited by them. There may be
> some properties that are universally restricted to subjects [fn. omitted], but
> there are certainly none that they always have. Rather, we find properties that
> are exhibited by subjects in a wide range of languages, and which may be
> plausibly argued to be restricted to subjects in some of them.

This observation is not surprising – it is in line with the way typological properties typically apply (Comrie 1989). However, it violates the usual formalist preference for absolute universals, and thus is an important caveat for any formally based theory of subjects. In addition, the fact that typological properties typically emerge as tendencies rather than absolutes is itself something that needs to be explained.

The first property is that if a verb has an Agent argument, the Agent is realized as subject.

(1)   a. Predicate: 'eat'; Agent: 'the kid'; Patient: 'the sandwich'
          b. The kid ate the sandwich
          c. *The sandwich ate´ the kid.

A verb like the putative *eat´* in (1c), in which the Patient is realized syntactically as subject and the Agent as object, is disallowed. Of course, while all Agent arguments are subjects, not all subjects are Agents. If the verb does not have an Agent argument, the subject will express some other thematic role. A special case of this is the passive construction, in which the Agent loses its expressed-argument status (Chomsky 1981, Bresnan 2001).

Another property of subjects is that the addressee of an imperative is a subject. This can be seen in each of the following imperatives: the addressee can have a variety of thematic roles, not necessarily Agent, but it must have the syntactic status of subject.

(2)   a. Eat the sandwich!
          b. Go to school!
          c. Freeze, if that's what you want! (Parent to child who refuses to put on a
             coat in freezing weather)
          d. Be happy!
          e. Be arrested by the municipal police, not the state police!

Another property which is apparent in the English imperative examples, although more clearly in other languages, is that the subject is more susceptible to being realized as a covert (null or empty) pronoun. It is telling that the empty-pronoun construction (or pro-drop) is often referred to in the

theoretical literature as the null-subject construction, a name which is based on this higher susceptibility. We will discuss the facts in more detail in Chapter 2.

A frequently discussed property of subjects is anaphoric prominence. The exact details vary from language to language (as will be discussed in Chapter 2), but one clear consequence which can be seen in all languages with reflexive pronouns[1] is that, in a transitive clause in which the subject and object are coreferential, it is the subject which is expressed as a full NP and the object as the reflexive pronoun.

(3)   a. Pnina saw herself.
        b. *Herself saw Pnina.

In some languages the antecedent of a reflexive must be a subject, while in others (like English) it just has to have higher prominence, but in either case the most prominent element of the clause is the subject.

An anaphoric construction which does not exist in English, but in which the greater prominence of the subject is again apparent, is the switch-reference construction, in which a clause marks the anaphoric relation (coreference or disjoint reference) between its own subject and the subject of a superordinate and/or coordinate clause. This is exemplified in the following Diyari sentences (Austin 1981).

(4)   a. Karna wapa- rna   warrayi, jukudu   nanda- lha.
           man   go-   PART AUX   kangaroo kill-   IMPLIC.SAME
           'The man went to kill a kangaroo.'

        b. Karna- li   marda matha- rna   warrayi, thalara
           man-   ERG stone bite-   PART AUX   rain

           kurda- rnanthu.
           fall-   IMPLIC.DIFF
           'The man bit the stone so the rain would fall.'

In (4a), the clauses have coreferential subjects, so the "same" morpheme is used in the subordinate clause. In (4b), on the other hand, the subjects are disjoint in reference, and the "different" morpheme is used.

---

1 An anonymous reviewer suggests that the data from Samoan in Chapin (1970) may be a counterexample. Chapin observes that there is no subject/non-subject asymmetry for a pronoun with a reflexive interpretation; the only condition is that the antecedent must precede the pronoun. However, he also notes that there are no distinct reflexive pronouns in the language. Since the Samoan forms are simply undifferentiated anaphoric elements, there is no reason to expect a restriction to subject.