FACTORS GOVERNING THE MORPHOLOGICAL CODING OF REFERENTS IN HAUSA NARRATIVE DISCOURSE

JAGGAR, P. 1985

Published on demand by

UNIVERSITY MICROFILMS INTERNATIONAL

Ann Arbor, Michigan, U.S.A. · London, England

This is an authorized facsimile
printed by microfilm/xerography on acid-free paper
in 1985 by
UNIVERSITY MICROFILMS INTERNATIONAL
Ann Arbor, Michigan, U.S.A.

1416

8513129

Jaggar, Philip John

FACTORS GOVERNING THE MORPHOLOGICAL CODING OF REFERENTS IN HAUSA NARRATIVE DISCOURSE

University of California, Los Angeles PH.D. 1985

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International

300 N. Zeeb Road
Ann Arbor, MI 48106

UNIVERSITY OF CALIFORNIA

Los Angeles

Factors Governing the Morphological Coding of Referents in Hausa Narrative Discourse

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Linguistics

by

Philip John Jaggar

1985

The dissertation of Philip John Jaggar is approved.

John W. Du Bois

Robert S. Kirsner

Pamela Munro

Thomas G. Penchoen

Paul Schachter

Russell G. Schuh, Committee Chair

University of California, Los Angeles

1985

TABLE OF CONTENTS

	Page
List of Abbreviations	vi
List of Tables and Figures	viii
Acknowledgements	xi
Vita and Publications	xii
Abstract	xiv
CHAPTER I, Introduction	1
1. General aims and scope of the study	1
2. Data base	5
CHAPTER II, Strategies for Marking First-Mention Indefinite Nominals	7
1. Introduction	7
1.1. Previous Characterizations of the Indefinite Specifier	8
2. The Hypothesis	10
2.1. Categories of Indefinite NP which Generally Resist IS-Marking	13
3. Empirical Validation of the Hypothesis	26
4. First-Mention Human Indefinites	31
4.1. Subject Human Indefinites [+IS]	32
4.2. Subject Human Indefinites [-IS]	35
4.3. Nonsubject Human Indefinites [+IS]	38
4.4. Nonsubject Human Indefinites [-IS]	41
5. First-Mention Animal Indefinites	49
5.1. Subject Animal Indefinites [+IS]	50
5.2. Subject Animal Indefinites [-IS]	51
5.3. Nonsubject Animal Indefinites [+IS]	54
5.4. Nonsubject Animal Indefinites [-IS]	55
6. First-Mention Inanimate Indefinites	57
6.1. Subject Inanimate Indefinites [+IS]	58
6.2. Subject Inanimate Indefinites [-IS]	59
6.3. Nonsubject Inanimate Indefinites [+IS]	60
6.4. Nonsubject Inanimate Indefinites [-IS]	61
7. The Disjunctive 'Other' Reading of the IS	63
8. Summary	65
Notes to Chapter II	68

CHAPTER III, Conditions on the Selection of Zero Anaphora and Pronominal Anaphoric Reference 74

1. Introduction 74
2. Data Base and Methodology 75
3. Distribution of the Three Referential Classes: an Overview 77
3.1. Zero Anaphoric Reference 84
3.2. Pronominal Anaphora 85
3.3. Nominal Coreference 86
4. Anaphoric Reference to the Direct Object. . 87
4.1. Direct Object Pronominal Anaphora 93
4.1.1. Human Referents [+ DO PRO] 93
4.1.2. Animal Referents [+DO PRO] 95
4.1.3. Inanimate Referents [+DO PRO] 96
4.2. Direct Object Zero Anaphora 98
4.2.1. Human Referents [+DO ZERO] 98
4.2.2. Animal Referents [+DO ZERO] 101
4.2.3. Inanimate Referents [+DO ZERO] 103
5. Nonoccurrence of Material Associated with Clausal Subjects 106
5.1. Zero Anaphora of the Subject Argument . . 107
5.2. Subject Concord-Drop 114
5.3. Auxiliary-Drop 117
5.3.1. Stative Auxiliary-Drop 118
5.3.2. Purpose and Participial Clauses 120
6. Summary 122
Notes to Chapter III 124

CHAPTER IV, Charting and Retrieving Referents with Lexical NP Configurations 128

1. Introduction 128
2. The Distributional Data on NP Categories: a Quantitative Overview 130
3. The Hypothesis 132
4. Empirical Verification of the Claims . . . 134
4.1. Ease/Difficulty of Referent-Accessibility and Morphological Coding 135
4.1.1. Noun + Demonstrative 144
4.1.2. Noun + Determiner 149
4.1.3. Simple Nominal 153
4.1.4. Proper Name 155
4.1.5. Noun + Possessor 157
4.2. Coreferential NP Choices Not Determined by Text Discontinuity 158
4.2.1. New Referent-Anchoring 160
4.2.2. Coding of Culturally-Prominent Referents 168
5. Summary 171
Notes to Chapter IV 173

CHAPTER V, Referential Choice, Subject-Switching/-Preservation and Larger Discourse Breaks . 179

1. Introduction 179
2. Subject-Preservation and Subject-Switching: a Quantitative Overview 181
3. Subject-Preservation and Referential Choice 185
3.1. Subject-Ellipsis 185
3.2. Nominal Coreference 188
4. Subject-Switching and Referential Choice. . 191
4.1. Subject-Ellipsis 191
4.2. Nominal Coreference 196
5. Paragraph-Signaled Discourse Breaks and Referential Choice 203
5.1. Subject-Ellipsis Within and Across P-Marked Boundaries 206
5.2. Subject Nominal Coreference Within and Across P-Marked Boundaries 215
5.3. Direct Speech Quotations and Referential Choice 230
6. Summary 239
Notes to Chapter V 241

CHAPTER VI, Concluding Remarks 246

Appendices 249

Bibliography 262

Published Data Sources 269

LIST OF ABBREVIATIONS

ASSOC	=	associative
ADV	=	adverb
CONC-Ø	=	concord marker-drop
COP	=	copula
DEM	=	demonstrative
DET	=	determiner
DIMIN	=	diminutive
DO	=	direct object
DO-Ø	=	direct object zero anaphora
EXIST	=	existential predicator
FUT	=	future
HABIT	=	habitual
IMPERS	=	impersonal
IMPFV	=	imperfective
IND PRO	=	independent pronoun
IO	=	indirect object
IS	=	indefinite specifier
LOC	=	locative
N	=	noun
NEG	=	negative
NP	=	noun phrase
PFV	=	perfective
PRESENT	=	presentative predicator

PRO	=	pronoun
REL	=	relative marker
REL FUT	=	relative future
STAT	=	stative
STAT-Ø	=	stative auxiliary-drop
SUBJ	=	subject
SUBJ-Ø	=	subject zero anaphora
SUBJNCTV	=	subjunctive
TOPIC	=	topicalized element
VN	=	verbonominal

LIST OF TABLES AND FIGURES

Table — Page

3.1. Summaries of Relative Frequencies of Referential Forms in Written and Spoken Narratives 78

3.2. Distribution of Direct Object Pronouns and Direct Object Zero Anaphora with respect to Time and Interference 92

3.3. Distribution of Various Categories of Subject-Related Omission with respect to Time and Interference 107

4.1. Relative Frequencies of Coreferential NP Categories in Written and Spoken Narratives 131

4.2. Distribution of NP Categories with respect to Time and Interference 137

4.3. Distribution in Written Narratives of NP Categories with respect to the Number of Clauses Present between Two Successive Mentions of a Referent 139

4.4. Distribution within Spoken Narratives of NP Categories with respect to the Number of Clauses Present between Two Successive Mentions of a Referent 140

4.5. Distribution within Written Narratives of NP Categories with respect to the Number of Potential Confusers Intervening Between Two Successive Mentions of a Referent . . 142

4.6. Distribution within Spoken Narratives of NP Categories with respect to the Number of Potential Confusers Intervening Between Two Successive Mentions of a Referent . . 143

4.7. The nan-Marked Demonstratives 145

Figure — Page

2.1. Written Narratives: IS-Marking of First-Mention Indefinites According to Semantic Class and (Initial) Syntactic Status 28

2.2. Average Number of Subsequent Mentions of Indefinites According to IS-Marking, Semantic Category and Syntactic Status . . 30

3.1. Distribution of the Three Referential Categories with respect to the Number of Intervening Clauses (Written Narratives) . 80

3.2. Distribution of the Three Referential Categories with respect to the Number of Intervening Clauses (Spoken Narratives) . 81

3.3. Distribution of the Three Referential Categories with respect to the Number of Intervening Confusers (Written Narratives) 82

3.4. Distribution of the Three Referential Categories with respect to the Number of Intervening Confusers (Spoken Narratives) 83

3.5. Distribution within Written Narratives of Direct Object Zero Anaphora and Direct Object Pronoun Anaphora with respect to Semantic Class 89

3.6. Distribution of Direct Object Pronouns and Direct Object Zero Anaphors in Written Narratives: Numbers and Percentages according to Semantic Class 90

5.1. Written Narrative Percentages of NP's and Ellipsis Selected for Preserving and Switching Subject Reference Within and Across Sentences 183

5.2. Oral Narrative Percentages of NP's and Ellipsis Selected for Preserving and Switching Subject Reference Within and Across Sentences 184

5.3. Written Narrative Percentages of Differential NP Configurations Chosen for Switching Subjects Within and Across Sentence Boundaries 201

5.4. Written Narrative Percentages of Nominal and Elliptical Subjects Selected where Prior Reference is Within or Across a Paragraph-Marked Boundary 205

5.5. Written Narrative Percentages of Differential NP Configurations Selected where Prior Mention is Within and Across Paragraph-Marked Boundaries 216

5.6. Written Narrative Percentages of Differential NP Configurations Selected to Encode Subject Arguments Immediately Following "Direct Speech" Quotations . . . 235

Acknowledgements

I would like to express my appreciation to the members of my Doctoral Committee for comments on earlier versions of this dissertation. Special thanks are due to John Du Bois, Robert S. Kirsner, Paul Schachter, and Russell G. Schuh, all of whom have maintained a keen and invaluable interest in all stages of the work. I am also indebted to Sani Ahmad Sufi and Haruna Waziri for providing some of the spoken Hausa texts. Finally, I must thank my wife, Gloria E. Hatrick, for proofreading and regular encouragement. None of the above should be held responsible for any of the remaining shortcomings.

VITA

June 14, 1945--Born, Bradford, England

1968--B.A., University of London

1973-76--Lecturer in Hausa, Department of Nigerian Languages, Bayero University College, Nigeria

1976-78--Lecturer in Hausa, Seminar für Afrikanische Sprachen und Kulturen, University of Hamburg

1978--M.Phil., University of London

1978-83--Teaching Associate, Department of Linguistics, University of California, Los Angeles

1981--M.A., University of California, Los Angeles

1981--Visiting Lecturer in Hausa, Department of African Languages and Literature, University of Wisconsin, Madison

1983--C.Phil., University of California, Los Angeles

1984--Lecturer in the Hausa Language, School of Oriental and African Studies, University of London

PUBLICATIONS

Jaggar, P.J.
1973a A Kano blacksmith's vocabulary. Kano Studies, New Series, 1:99-109.

1973b Kano City blacksmiths: precolonial distribution, structure and organisation. Savanna 2:11-25. Zaria, Nigeria: Ahmadu Bello University.

1976 Thematic emphasis in Hausa. Harsunan Nijeriya 6:63-75. Kano, Nigeria: Centre for the Study of Nigerian Languages, Bayero University College.

1977a Translation of the Secondary School Atlas for Kano, Kaduna and Sokoto States. Zaria, Nigeria: Northern Nigerian Publishing Company (Macmillan Ltd.).

1977b The nature and function of auxiliary verbs in Hausa. In (eds.) Paul Newman and Roxana Ma Newman, Papers in Chadic Linguistics, pp. 57-87. Leiden: Afrika-Studiecentrum.

1978 And what about...? -- Topicalisation in Hausa. Studies in African Linguistics 9:69-81.

1980 Review of Einführung in die Hausa-Sprache by Herrmann Jungraithmayr and Wilhelm J.G. Möhlig. Journal of African Languages and Linguistics 2:75-78.

1981a Review of Modern Hausa-English Dictionary by (eds.) Paul Newman and Roxana Ma Newman. Language 57:505-506.

1981b The Kano dialect of Hausa in the Polyglotta Africana. Afrika und Übersee 64:47-58.

1982a Varieties of passive in Hausa. In (ed.) William R. Leben, Papers from the Twelfth Annual Conference on African Linguistics. Studies in African Linguistics, Suppl. 8:73-77.

1982b Monoverbal imperative formations in Hausa: a striking case of analogic realignment. Journal of African Languages and Linguistics 4:133-56.

1983 Some dimensions of topic-NP continuity in Hausa narrative. In (ed.) Talmy Givón, Topic Continuity in Discourse: Quantitative Cross-Language Studies (Typological Studies in Language, vol. 3), pp. 365-424. Amsterdam: John Benjamins.

ABSTRACT OF THE DISSERTATION

Factors Governing the Morphological Coding of Referents
in Hausa Narrative Discourse

by

Philip John Jaggar
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 1985
Professor Russell G. Schuh, Chair

This dissertation is a discourse-based investigation of a range of referential types in Hausa (Chadic, Afro-asiatic) narrative. In dealing with the cognitive task of introducing and then tracking referents in narrative discourse, a speaker/writer is constantly faced with the problem of selecting the referential form most appropriate in a given context, and the basic question I attempt to answer is: What are the factors which condition the ongoing choices we encounter? Whilst the study itself is primarily descriptive, the findings are viewed as having implications for general theoretical approaches to the problem of referentiality.

The data are taken mainly from the written stories of

Imam (1970, 1971), supplemented, for purposes of comparison, by several spoken narratives. In order to determine the frequency with which the various linguistic forms occur, I make use of statistical, cross-text generalizations.

The main hypothesis states that, in certain areas of Hausa grammar, the choice between competing referential options is influenced by the "discourse salience/deployability" of a referent, defined as the degree to which it is deployed as an autonomous, pivotal argument within the discourse. The more prominent the discourse role of a referent, the heavier the morphological coding it receives; conversely, the less salient the referent, the more attenuate the marking. Humans are generally the most salient arguments, and are assigned heavier coding than nonhuman (animal and inanimate) entities. This interpredictability is manifested in two domains: indefinite reference, where human participants carry an "Indefinite Specifier" with much greater frequency than first-mention nonhumans; definite reference in the postverbal direct object position, where human arguments favor coding with a full pronoun, and less salient nonhumans tend to be marked with zero anaphora.

Finally, there are choices which seem to be determined by more general constraints on short-term memory. These include omission of material associated with the

preverbal subject argument, and the competition between a simple noun and a noun plus deictic.

CHAPTER 1

INTRODUCTION

1. General aims and scope of the study

The recent surge of interest in discourse analysis has led to increased interest in the study of anaphoric reference and the relationship between the discourse structures being presented, e.g. narratives, expository texts, conversational exchanges, and the choice of referential forms, e.g. overt lexical reference as opposed to less explicit strategies such as pronouns and zero anaphors. This approach stems from the increasingly accepted belief that the correct interpretation of anaphoric relations cannot be stated simply in terms of the syntactic or semantic properties of individual clauses, and some of the more prominent studies devoted to these and other related issues include Chafe (1980b), García (1975), Givón (1979a, 1983), Grimes (1975, 1978), Hinds (1977, 1978b), and Klein-Andreu (1983a). My general sympathies lie, therefore, with the position that the function and meaning of referential types can only be sensibly characterized through scrutiny of their formal distribution and behavior in coherent, multisentential chunks of discourse. Lan-

guage being functional, the use of addressee- and/or addressor-oriented strategies is explicable in terms of the communicative aspects of language, rooted in discourse. This dissertation thus proceeds "...on the assumption that actual language use is PURPOSEFUL -- that is, that users (speakers/writers) employ linguistic elements so as to contribute to a particular communicative intent, which should be reflected in the contexts in which the elements occur" (Klein-Andreu 1983a:xvi), and investigates the domains of indefinite and definite reference in Hausa, an SVO language belonging to the Chadic (Afroasiatic) language family.

The only published, discourse-based attempt at "explaining" the distribution of the various reference-types encountered in natural discourse is the one presented in Jaggar (1983). Although this study has the merit of covering a quite comprehensive range of reference-types, it suffers from two major drawbacks: firstly, its wide sweep has meant that only a superficial investigation of the various coding-points was possible; and secondly, the narrow analytical methodology had the effect of excluding, from consideration, a number of important referential phenomena, some of which are in fact damaging to the central claims underpinning Givón's (1983) volume. The basic aim of this dissertation, therefore, is to provide a necessary and adequate elaboration of this treatment.

The heart of the dissertation is divided into four main chapters. Chapter 2 investigates the devices exploited to mark referents when they are first introduced into a text. Two options are available, entailing the use, or nonuse, of what I term the **"Indefinite Specifier"**. I show that the acceptability-conditions on selection of the Indefinite Specifier to mark a first-mention indefinite relate to the **"Discourse-Deployability/Salience"** of an indefinite nominal, i.e. its potential for deployment as an ongoing, pivotal argument in the ensuing discourse, and measured in terms of the number of subsequent mentions -- the so-called **"Salience:Coding"** hypothesis. I demonstrate further that use of the Indefinite Specifier correlates with two factors: the semantic class membership of the indefinite, and its syntactic role on first-mention, with highly-salient, highly-codable human subjects attracting by far the highest ratio of Indefinite Specifiers on first mention.

Chapter 3 moves on to the domain of definite reference and utilizes the same salience:coding hypothesis to account for the cross-text distribution of the two major classes of nonexplicit reference -- pronominal and zero anaphora. The analysis reveals that, in the postverbal direct object position, the choice of a full pronoun in preference to a zero anaphor is again influenced by the discourse salience of the referent, with humans consist-

ently favoring pronouns, and nonhuman arguments (animals and inanimates) assigned a greater proportion of zero anaphors. Chapter 3 also looks at the contexts in which zero anaphoric reference to the preverbal subject argument is permissible.

In Chapters 4 and 5, we move on to consideration of the domain of nominal coreference, i.e. use of a noun, with or without accompanying modifiers, to refer to an identifiable, definite referent. Chapter 4 takes a close look at a finely-graded continuum of lexical alternatives, including simple noun, noun with determiner, and noun plus demonstrative, and attempts to determine some of the factors which condition selection of these referential types. I show that although cognitive constraints do play an important part in the choice of a coreferential NP, i.e. the encoder (speaker/writer) considers that the decoder (hearer/reader) requires lexical assistance to retrieve a referent, there are other factors which can prompt the decoder to exploit NP coreference.

Finally, Chapter 5 investigates the relationship between referential choice and three phenomena -- subject-switching, subject-maintenance, and the intervention of larger discourse breaks such as episode-boundaries.

2. Data base

The bulk of the textual material for this dissertation is taken from the narrative, non-speech, portions of the fictional stories in Imam (1970, 1971). For purposes of comparison, I also collected four spontaneously-produced oral narratives -- two personal stories, and two "Pear Film" accounts. One of the personal narratives was related by an elderly man, and contained his recollections of the arrival of the British colonial forces in northern Nigeria; and the other was a young man's description of a "brush-with-death" type experience. The so-called "Pear Film" was designed for viewing by speakers of different languages, who would afterwards relate their perceptions of the film (for an extensive discussion, see Du Bois 1980a). I showed the Pear Film to two young male Hausa speakers resident in the Los Angeles area, and then tape-recorded the narratives they produced.

My decision to use narrative texts as the baseline for this study should not be construed as implying that I consider narrative discourse to have any primacy over other discourse-types, e.g. spontaneous conversation, expository prose etc. (cf. Tannen 1982), a position in fact taken by such writers as Givón (1979b:52) and Hopper (1979). Rather, it is dictated by the following considerations: it is relatively easy to keep track of referential

information in such compact, integrated texts; narratives tend to have easily recognizable internal boundaries; and the written narrative corpus (Imam 1970, 1971) contains a useful cross-section of semantic classes and referential forms, including animal as well as human and inanimate entities, and proper names in addition to common nouns.

CHAPTER 2

STRATEGIES FOR MARKING FIRST-MENTION INDEFINITE NOMINALS

1. Introduction

In this opening chapter I investigate the referential strategies available for encoding the introduction of indefinite arguments into Hausa narrative discourse -- humans, animals and inanimates which are nonidentifiable in the sense that the "encoder", i.e. speaker/writer, presumes the "decoder", i.e. hearer/reader, unable to establish any direct and unique link between the indefinite form used and the real-world referent it is intended to denote (Chafe 1976; Du Bois 1980b).

The exponent of initial-mention indefinite reference in Hausa is either zero, i.e. a simple nominal without any morphological trappings, or a nominal preceded by a pragmatic operator I shall henceforth refer to as the **"Indefinite Specifier"** or "IS", the occurring, gender/number-sensitive forms of which are **wani** (m.sg.)/**wata** (f.sg.)/**wa(ɗan)su** (pl.).[1] Sentences (1a-c) illustrate the use of this morpheme in a typical environment -- to modify human indefinites occupying the preverbal "subject" slot

on first mention:[2]

(1) a. Wani yaro ya zo.
IS boy-SUBJ he-PFV come
'A (certain) boy came.'

b. Wata yarinya ta zo.
IS girl-SUBJ she-PFV come
'A (certain) girl came.'

c. Wa(d̂an)su yara suka zo.
IS children-SUBJ they-PFV come
'Some children came.'

The basic purpose of the present chapter is to achieve some generalization in determining the discourse-pragmatic and semantic factors which condition selection of one or the other of these two referential options for first-mention indefinites -- bare nominal or IS + nominal.

1.1. Previous characterizations of the Indefinite Specifier

Pedagogical grammars of Hausa have little to offer on the semantics and distribution of the IS. Kraft and Kirk-Greene (1973:54), for example, write, somewhat enigmatically, that it should be used "When the sense requires a certain...", e.g. wani yaro ya zo 'a certain boy has come'. Cowan and Schuh (1976:151-52) provide a rather ominous list of glosses for what they term the "Indefinite Pronoun or Adjective", including 'a(n)/some, a certain; someone, (a) certain one/ones; another, some more,

other(s)', and offer the encoder-oriented definition that "what all these meanings have in common is indefiniteness or not being specifically known to the speaker." Additional evidence of the complex semantics of the IS comes from the fact that in his dictionary entry for this formative, Abraham (1962:922-23) lists the following five separate glosses: 'some/someone; another (person); a certain (person); one...the other; any/anyone'.

The two published attempts at "explaining" the various phenomena -- semantic, syntactic, and stylistic -- associated with the IS appear in Skinner (1974) and Jaggar (1983). Neither study, however, does justice to what is an extremely subtle and elusive referential system. Skinner (1974:252-53), for instance, proposes that the Indefinite Specifier is exploited to encode referents which are [+particular], [-already mentioned or understood], and [+important], and he links nonoccurrence of an IS to the absence of one or both of the two features [+particular] and [+important]. Beyond saying that the IS serves to mark a newly-introduced, often episode-initial "particular, important character (or, much more rarely, thing)" (1974:253), Skinner does not address the obvious question of how the use or nonuse of the IS might correlate with the lexicosemantic class or syntactic status of a first-mention indefinite NP. In Jaggar (1983:399ff.), I used the term "Referential-Indefinite Marker" (RIM), and

tried to demonstrate, via statistically-based generalizations, that such correlations do exist, concluding that "[+human] arguments, whatever their grammatical function within the clause, which are REF-INDEF according to the above criteria are, almost without exception, modified by a pre-positional RIM", and that "regarding first mention of inanimate, non-locative entities, perusal of any Hausa text will indicate that the decision to employ an RIM or not can go either way basically" (1983:401). Whilst both proposals are not without their merits -- some of the respective claims in fact overlap -- closer scrutiny of the formal distribution of the IS shows them both to be oversimplifications, and the remainder of this chapter attempts to provide a necessary refinement of these preliminary characterizations.

2. The hypothesis

The essential meaning signaled by the Indefinite Specifier, and underlying all its occurrences, is that is serves to particularize/specify/categorize/individuate etc. a first-mention indefinite, singling it out from the wider category of which it is a type-token, and thereby focusing the decoder's attention on the referent so-marked. Given this approximate semantic characterization, the question arises "Why should the encoder wish to direct the

decoder's attention to a particular first-mention indefinite?". In answer to this obvious but important question, I shall argue for the explanatory and predictive value of the following claims:

(2) A crucial manifestation of this core semantic substance is to be found in the context of integrated discourse, where there is a discernible correlation between the **"discourse-deployability"** or **"discourse-salience"** of a first-mention indefinite -- defined as the degree to which a given referent is deployed as an autonomous, ongoing, pivotal argument within the discourse, serving as a potential basis for further discourse mentions -- and the assignment or not of an accompanying Indefinite Specifier.[3] Selection of an IS is thus seen as prompting the decoder to establish a "new cognitive file" on the referent so-marked (Du Bois 1980b:220-21), creating the expectation that additional information will be added.[4] The basic prediction states that the more salient or prominent the discourse role of an argument -- measured in terms of the number of subsequent discourse mentions (cf. Figure 2.2) -- the greater the likelihood that it will receive an IS on first appearance. And conversely, we anticipate that a relatively low degree of deployability will correlate with a lower

average incidence of IS-marking. I hope to demonstrate further that selection of the IS-marking option is sensitive to two interlocking variables:

(a) The lexicosemantic class of the indefinite NP. In this regard, human participants generally represent the most salient, discourse-deployable entities and so are brought into prominence via IS-marking more consistently than their nonhuman (animal and inanimate) counterparts.

(b) The syntactic status of an NP on initial mention. Referents introduced in the prominent role of clausal subject display a higher rate of discourse-deployability than do nonsubject indefinites, and as a consequence receive a greater proportion of IS's.

Before proceeding to a quantitative validation of the foregoing claims, it is important to note that there are a number of environments in which the use or nonuse of an IS allows some observations which are relevant to the hypothesis. Illustrating these categories at this stage has the additional advantage of introducing the reader to some of the acceptability conditions on IS-marking of indefinites, and also serves to highlight some of the primary discourse functions of the IS.

2.1. Categories of indefinite NP which generally resist IS-marking

As Russell Schuh (p.c.) has correctly reminded me, the use or nonuse of an Indefinite Specifier on a first-mention indefinite frequently correlates with the semantic distinction between, respectively, specific and nonspecific generic reference. In other words, an important function of the IS-marking strategy, armed with its [+particular] semantics, is its exploitation in forcing a specific interpretation on its referent, thereby blocking the possibility of a nonspecific generic reading which the absence of an IS, i.e. a bare stem, would favor. Once the IS has been used to particularize an initial-mention indefinite, separating it both from other members of its class and also preventing identification with the genus itself, it is invested with the potential for further reference, i.e. it becomes discourse-deployable in my terms, though as Du Bois (1980b:209) observes, this potential for deployment is not always utilized. And conversely, there will be less of a tendency to allow [-IS] generic forms to control subsequent anaphoric (pronoun or zero) and coreferential (NP) mentions, since no individual autonomous identity is specified. Fragment (3) nicely illustrates some of the above correlations:[5]

(3) a. Ø Suka gamu da wani mutum
SUBJ-Ø they-PFV meet with IS man-ASSOC

b. Ø ya-na tafiya shi kaɗai...
SUBJ-Ø he-IMPFV travel-VN he-IND PRO only
c. Da mutum ya gan shi
as soon as man-SUBJ he-PFV see him-DO PRO
d. Ø ya ga ɓarawo kuntukurmi.
SUBJ-Ø he-PFV see thief outright-DO
e. Da Ø suka gamu da mutumi-n nan
when SUBJ-Ø they-PFV meet with man-DET DEM ASSOC
f. sai Ø ya wuce...
then SUBJ-Ø he-PFV pass by

(personal notes)

'They met a (certain) man traveling by himself...As soon as someone/anyone saw him, he would have recognized an outright thief. When they met this man he passed by...'

In (3a), the first-mention indefinite associative human NP wani mutum 'a (certain) man' is typically supported by an IS -- see 4.3 below for details -- and is immediately picked up with a direct object pronoun shi 'him' in (3c) (cf. Chapter 3), a coreferential associative nominal + demonstrative strategy mutumin nan 'this man' in (3e) (cf. Chapter 4), and again with subject agreement on the auxiliary ya in (3f) (cf. Chapter 3). In the intervening clause (3c), however, we encounter an example of the same nominal -- mutum 'someone, anyone (man)' -- this time without any accessory IS because a nonspecific generic interpretation is intended. Excerpt (4) also illustrates several important principles:

(4) a. ...Ø ya gano wani wuri ne
SUBJ-Ø he-PFV see IS place-DO COP
b. inda ɓarayi suka ɓoye wadansu kuɗi
where thieves-SUBJ they-PFV hide IS money-DO

gangariya-r-su ciki-n wani kogo-n dutse.
pile-of-them inside-of IS hollow-of rock
c. Sambo ya-na tsammani kuɗi-n sarki-n
Sambo-SUBJ he-IMPFV think money-of emir-of
gari-n ne,
town-DET COP
d. don yau bai fi kwana uku ba
because today it-NEG-PFV exceed day three NEG
e. da wadansu ɓarayi suka shiga
that IS thieves-SUBJ they-PFV enter
taska-r-sa,
store-of-his-LOC
f. Ø suka yi masa ƙat,
SUBJ-Ø they-PFV do to him-IO PRO clean out
g. a-na kuwa tsakiya-r nema-n-su ne...[6]
IMPERS-IMPFV and middle-of look-of-them-VN COP

(Imam 1970:116-18)

'...he (Sambo) saw a place where thieves had hidden some piles of cash in a cave. Sambo thought it was the money of the emir of the town, because only three days previously some thieves had broken into his store-room and cleaned him out, and they were being looked for...'

Fragment (4) contains a liberal sprinkling of IS-marked indefinite NP's. Our immediate concern is with the differential marking used to encode first-mention reference to the two separate sets of 'thieves' -- the bare nominal subject [-IS] ɓarayi 'thieves' in (4b) vs. the subject NP [+IS] wadansu ɓarayi 'some thieves' in (4e). The two indefinites differ in the following manner: the bare stem ɓarayi is intended as a nonspecific generic mention, with consequent neutralization of the indentifiable:nonidentifiable contrast, hence the absence of any IS; in contrast, the IS-marked NP wadansu ɓarayi is so coded because the writer means this particular referent to be interpreted as

a specific indefinite. The two forms also differ somewhat in their discourse prominence: the [-IS] generic (4b) mention "decays" immediately, i.e. it controls no further discourse appearances, and the [+IS] specific (4e) referent has three subsequent mentions, including subject auxiliary-agreement on suka in (4f), and a possessive pronoun suffix /-su/ in (4g). Extract (4) also contains three inanimate referents which bear an IS on first mention -- wani wuri 'a place' in (4a), and wadansu kudi 'some money' and wani kogon dutse 'a cave' in (4b). The IS-marking of the two locative expressions can be accounted for by the criterial [+particular] feature I suggested for the IS -- spatial (and temporal) shifts are important elements in a continuous narrative -- even though they tend to be mentioned only once -- and Hausa uses the IS to formally express this strongly (re)orientational role.[7] And the occurrence of the IS-marked direct object indefinite wadansu kudi gangariyarsu 'some piles of cash' in (4b) is explicable in terms of its discourse-deployability -- it is particularized because it is destined to act as an important "prop" until the very end of the story.

Before proceeding any further I would like to point out a possible difficulty concerning the appropriate interpretation of certain first-mention indefinites not modified with a preposed IS. I encountered some cases where it proved difficult to decide whether the decision

not to use an IS, i.e. leaving the simple noun stem, was intended to signal a nonspecific generic meaning or a specific interpretation for the referent -- cf. the potential ambiguity arising in English from use of the "zero-form" plural indefinite in 'I saw men/apples' noted in Wald (1983:93). Because Hausa seems to "undercode" this particular semantic distinction, I think it reasonable, in some cases at least, to allow for the possibility that nonselection of the IS could lend itself to either interpretation -- cf., for example, mutum 'a man' in (5b) below.[8] Whatever interpretation one chooses to place upon this type of referential form, it remains the case that use of a bare nominal to code indefinite reference normally acts as a general constraint against independent deployability, in contrast, that is, to IS-marking which provides the indefinite with the potential for autonomous deployability. As we see in due course, the vast majority of bare stem tokens come from the class of inanimate indefinites (cf. Figure 2.1), for these entities are not usually invested with the high degree of discourse salience which is characteristic of humans for instance (cf. Figure 2.2), and so are not IS-marked with the same frequency. My prime concern, however, is with first-mention indefinites which are unambiguously specific in reference but not modified with any IS, and how the distribution of these [-IS] forms might differ from that of their [+IS]

counterparts.

In keeping with strong cross-language tendencies (Du Bois 1980b:209ff.), indefinite predicate nominals tend to exhibit a reduced capacity for discourse-deployment and so do not, as a rule, carry any IS. In (5), for instance:

(5) a. Da gari ya waye
when dawn-SUBJ it-PFV break
b. sai sarki ya ga mutum bisa rufi-n
then emir-SUBJ he-PFV see man-DO on roof-of
soro-n-sa.
house-of-his
c. Ga shi tsamo-tsamo ciki-n gida,
PRESENT he-DO PRO dripping wet inside-of house
d. sarki ya yi tsammani mahaukaci ne.
emir-SUBJ he-PFV do thinking madman COP

(Imam 1970:148)

'At daybreak the emir saw a man on the roof of his house. There he was, dripping wet in the palace, and the emir thought he was a madman.'

the categorizing predicate nominal mahaukaci 'madman' in clause (5d) is nondeployable and so incompatible with IS-marking. The writer is not referring to a specific, concrete, individuated NP 'madman', and does not plan to use the NP as an autonomous discourse-participant controlling future anaphoric reference; nor does Imam go on to discuss any actions or events which are in any sense associated with the conceptual frame denoted by the same NP. Rather it is the general class properties of the nominal which are abstracted away and used to categorize the referent in question, hence the lack of any modifying

IS. Fragment (6) exemplifies the same phenomenon in a simple, non-embedded copular construction:

(6) a. To, ashe shi falke ne,
OK really he-IND PRO-SUBJ trader COP
b. ba inda ø ba ya buga-wa...
NEG EXIST where SUBJ-ø NEG he-IMPFV hit-VN

(Imam 1970:174)

'In fact he was a long-distance trader, traveling everywhere...'

Again, it is the the quality designated by the predicate nominal falke 'long-distance trader' in (6a) which is exploited, not its potential for concrete interpretation. And in (7):

(7) a. ø Ya-na nan
SUBJ-ø he-IMPFV there-LOC
b. ø ya-na jifa-n doki da dutse
SUBJ-ø he-IMPFV throw-of-VN horse with stone
a banza,
in vain-ADV
c. ø ya-na ihu kama-r mahaukaci...
SUBJ-ø he-IMPFV shout like-of madman

(Imam 1970:150)

'There he was, vainly throwing stones at the horse, and shouting like a madman...'

the NP mahaukaci 'a madman' in the (7c) comparative "like" phrase is similarly nondeployable -- it is merely the abstract quality of being a 'madman' which is being expressed in the predication of 'like a madman', not any concrete, tangible entity. No attempt is made to subsequently track the identity of this NP in the discourse.'

It is of interest to note, however, that predicate indefinites may become eligible for IS-marking in contexts where some further discussion or elaboration of the predicate nominal is planned.[9] Example (8) illustrates this phenomenon:

(8) a. Waɗannan 'Yan Gadi-n waɗansu sakarkaru-n
DEM guards-DET-SUBJ IS idiots-of
ƙauyawa ne,
villagers COP
b. sabo-n shiga.
new-of enter-VN

(Imam 1971:21)

'These guards were (certain) idiot villagers, newly arrived.'

In (8a), the categorizing indefinite NP waɗansu sakarkarun ƙauyawa '(certain) idiot villagers' is particularized with a preceding IS waɗansu, not because it is destined to be anaphorically picked up as such, but because its mention provides an important rationale or motivation for many of the (foolhardy) deeds which are later perpetrated by the same 'guards', and which form the event backbone of the plot.[10] Whilst it is true, as in excerpts (5-7), that it is the general class properties of the NP in question which are are again being abstracted away, I would suggest that the IS is used here in order to highlight these particular character flaws and so draw the decoder's attention to them. This may be compared with the example cited in (5d), for instance, where IS-marking would be

less natural.[11]

Extract (9) is also of some interest in this regard:

(9) a. Kalala Allah ya yi shi
Kalala-TOPIC Allah-SUBJ he-PFV make him-DO PRO
wani iri-n mutumi-n da ne,
IS kind-of man-of before COP
b. ba abi-n da ø ya fi
NEG EXIST thing-DET REL SUBJ-ø he-PFV exceed
da ƙi
with dislike-VN
c. illa a yi masa abinci
except IMPERS-SUBJNCTV make for him-IO PRO food-DO
d. ø ya ci shi kaɗai.
SUBJ-ø he-SUBJNCTV eat he-IND PRO only

(Imam 1970:15)

'(And) Kalala, Allah had made him a particular type of old-fashioned fellow, he disliked nothing more than for food to be prepared for him and he eat alone.'

In (9a) we encounter another instance of an IS-marked categorizing predicate nominal -- wani irin mutumin da 'a certain type of old-fashioned fellow'. In the environment of an NP modified by the preposed sortal nominal iri 'kind, sort' plus a genitival linker /-n/, e.g. irin...X 'kind/sort of...X', the presence of an IS before the type-NP irin is in fact obligatory if an indefinite interpretation is intended -- cf. Halima ta sami iri-n mutumi-n da ø ta-ke so (Halima-SUBJ she-PFV find kind-of man-DET REL SUBJ-ø she-IMPFV want-VN) 'Halima found the kind of man she wanted', where nonoccurrence of the IS forces a definite reading on the NP. I believe too that its co-occurrence with the indefinite sortal nominal provides indepen-

dent support for the basic semantic characterization I have proposed for the IS, i.e. it conveys a particularizing/individuating/categorizing etc. force.

As noted by Givón (1977, 1979b:93ff., 1982:85ff.), NP's falling within the scope of negation are characteristically non-deployable ("non-referential" in his terminology, cf. fn. 3), and in Hausa such referents are, in the main, incompatible with IS-marking, for the obvious reason that since no existence is implied, they cannot be exploited to map autonomous, individuated arguments. In the following excerpt, the indefinites 'ya'ya 'children (10a)', ƙane 'younger brother' (10b), and wa 'older brother' (10c) are all non-deployable entities, occurring in negative (imperfective) associative environments, and carrying no IS's:

(10) a. ...don Ø ba shi da 'ya'ya,
because SUBJ-Ø NEG he-IMPFV with children-ASSOC
b. Ø ba shi da ƙane,
SUBJ-Ø NEG he-IMPFV with younger brother-ASSOC
c. Ø ba shi da wa.
SUBJ-Ø NEG he-IMPFV with older brother-ASSOC

(Imam 1971:1)

'...because he (the emir) didn't have children, he didn't have a younger brother, and he didn't have an older brother.'

Interestingly, Hausa does permit IS-coding of indefinite nominals under the scope of negation in certain circumstances.[12] Fragment (11) instantiates this possibility:

(11) a. To, duk ciki-n-su Ja'iru ba shi
OK all inside-of-them Ja'iru-SUBJ NEG he-IMPFV
da wani muhimmi-n ɗanuwa,
with IS important-of relative-ASSOC
b. uba-n-sa dai ya mutu,
father-of-his-SUBJ and he-PFV die
c. uwa-r-sa ta mutu,
mother-of-his-SUBJ she-PFV die
d. Ø ba shi da wa,
SUBJ-Ø NEG he-IMPFV with older brother-ASSOC
e. Ø ba shi da ƙane,
SUBJ-Ø NEG he-IMPFV with younger brother-ASSOC
f. Ø ba shi da ɗa,
SUBJ-Ø NEG he-IMPFV with son-ASSOC
g. Ø ba shi da jika...
SUBJ-Ø NEG he-IMPFV with grandson-ASSOC

(Imam 1970:127)

'Well, amongst them all, Ja'iru didn't have any particular/one/a single relative of importance, his father had died, his mother had died, he had no older brother, he had no younger brother, he had no son, he had no grandson...'

In (11b) the associative indefinite NP wani muhimmin ɗanuwa 'any particular/one/a single relative of importance' bears an IS despite the fact that it occurs in a negative-associative "have" predicate, and I would suggest that selection of the IS here functions to invest its referent with a kind of prominence, captured in the suggested English glosses 'any particular/one/a single'.[13] It would be perfectly possible to strip the IS wani away from its referent without impairing grammaticality in any way, but the putative nonuse of the IS would deprive it of any [+particular] value, leaving a more neutral reading.

Having established the relatively salient fact that the individual had no important kin to speak of, observe how, in the corresponding negative-associative clauses (11d-g), the writer then moves to the arguably more trivial task of breaking down this generalized statement, and reverts to the more regular convention of not marking negative predicate indefinites with any IS.

Fragment (12) exemplifies a first-mention inanimate indefinite -- wani amfani 'any particular/real value' in (12b) -- which is particularized via IS-marking, even though again it is located within a negative-associative construction:

(12) a. Da ø ya ga dai
when SUBJ-ø he-PFV realize and
b. kuka ba shi da wani amfani
crying-SUBJ NEG it-IMPFV with IS value-ASSOC
a iri-n wannan wuri,
in kind-of DEM place-LOC
c. sai ø ya tashi...
then SUBJ-ø he-PFV get up

(Imam 1970:91)

'When he realized that crying was of no particular/real value in that kind of place, he got up...'

A final category of first-mention indefinites which generally resist IS-marking concerns partially collocational verb + object conflations (Du Bois 1980b:214; Hopper and Thompson 1983). Example (13) is illustrative:

(13) ...sai wata rana da dare Kamaruzzaman
then IS day at night-ADV Kamaruzzaman-SUBJ
ya yi mafarki da uba-n-sa.
he-PFV do dream-DO with father-of-his-ASSOC

(Imam 1970:177)

'...then one night Kamaruzzaman had a dream about his father.'

The direct object NP mafarki 'dream' -- here conflated with the transitive verb yi 'do' to express the idea of 'do (have) a dream' -- is not being referred to as an autonomous, individuated entity. It is merely included in the unitary predicate concept of 'doing (having) a dream', and this lack of any real independent salience is communicated in the decision not to employ an IS. That said, it should be noted again, however, that examples are occasionally encountered where the encoder does exploit IS-marking to indicate that a particularizing interpretation is intended. In (14b), for instance:

(14) a. ø Ya-na nan
SUBJ-ø he-IMPFV there-LOC
b. sai ø ya tuna wata dabara
then SUBJ-ø he-PFV think of IS plan-DO
c. ø ya ce...
SUBJ-ø he-PFV say

(Imam 1971:48)

'There he was, when he thought of a (certain/particular) plan and said...'

the IS-marking of the direct object NP wata dabara 'a (certain/particular) plan' signals the writer's communi-

cative intent that this argument is to be considered [+particular]. Compare too such contrasting [±IS] sentence-pairs as Malamai suka ba shi (wata) shawara (teachers-SUBJ they-PFV give him-DO PRO (IS) advice-DO) 'The teachers gave him (a piece of/some particular) advice.'

The foregoing data have exemplified categories of indefinite reference where the selection or not of an IS has interesting semantic-pragmatic correlates, and they have been included in the cross-text counts which form the quantitative baseline for the remainder of this chapter (cf. Figures 2.1 and 2.2 below). A number of the examples cited have illustrated contexts in which exploitation of the IS merely serves to invest the indefinite referent with the semantic characteristics I have proposed, i.e. with a particularizing/individuating etc. force, and without necessarily preparing the nominal for further independent mention. As we shall now see, however, this particular function is the most important discourse-based manifestation of these semantics.

3. Empirical validation of the hypothesis

The remainder of this chapter attempts to confirm the claims advanced above in (2) that there is a direct and measureable correlation between the discourse-deployability of a first-mention indefinite and selection of an IS,

and that the intrinsic, perceptual salience of human subject arguments invests them with a relatively high degree of deployability and renders them the most natural candidates for IS-marking on initial mention. In order to permit text-based statistical generalizations at the level of discourse, I have isolated the first 700 or so tokens of initial-mention, third-person, indefinite reference to humans, animals and inanimates occurring in the narrative sections of Imam (1970, 1971). Figure 2.1 (cf. Appendix I) now supplies data on the distribution of IS-marking with respect to two variables: (1) the semantic class membership of each indefinite counted; (2) the syntactic slot in which each indefinite was first introduced into the text.

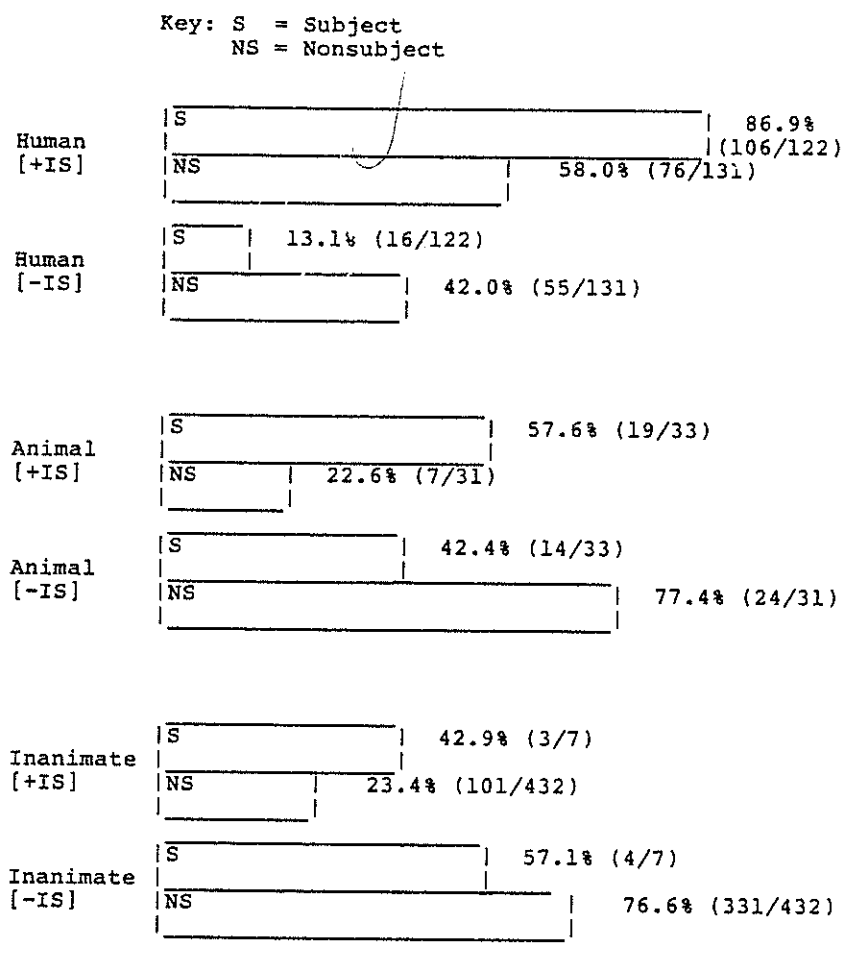


Figure 2.1. **Written narratives: IS-marking of first-mention indefinites according to semantic class and (initial) syntactic status.**[14]

The data summarized in Figure 2.1 essentially corroborate the claims. Thus:

(15) Human indefinites are assigned the highest percentage of IS's on first-mention, followed respectively by animal and inanimate indefinites.

(16) Subjects in general attract a greater proportion of IS's on first-mention than do arguments which are introduced in nonsubject roles.

Turning to the crucial matter of discourse-deployability, I suggested briefly in (2) that an intuitively reasonable way of objectifying this notion would be to calculate the **number of subsequent discourse-mentions** for each of the 700 or so indefinites counted in the corpus. Figure 2.2 (cf. Appendix II) now gives the numerical results of these counts, again with respect to the semantic class and syntactic role of each token.

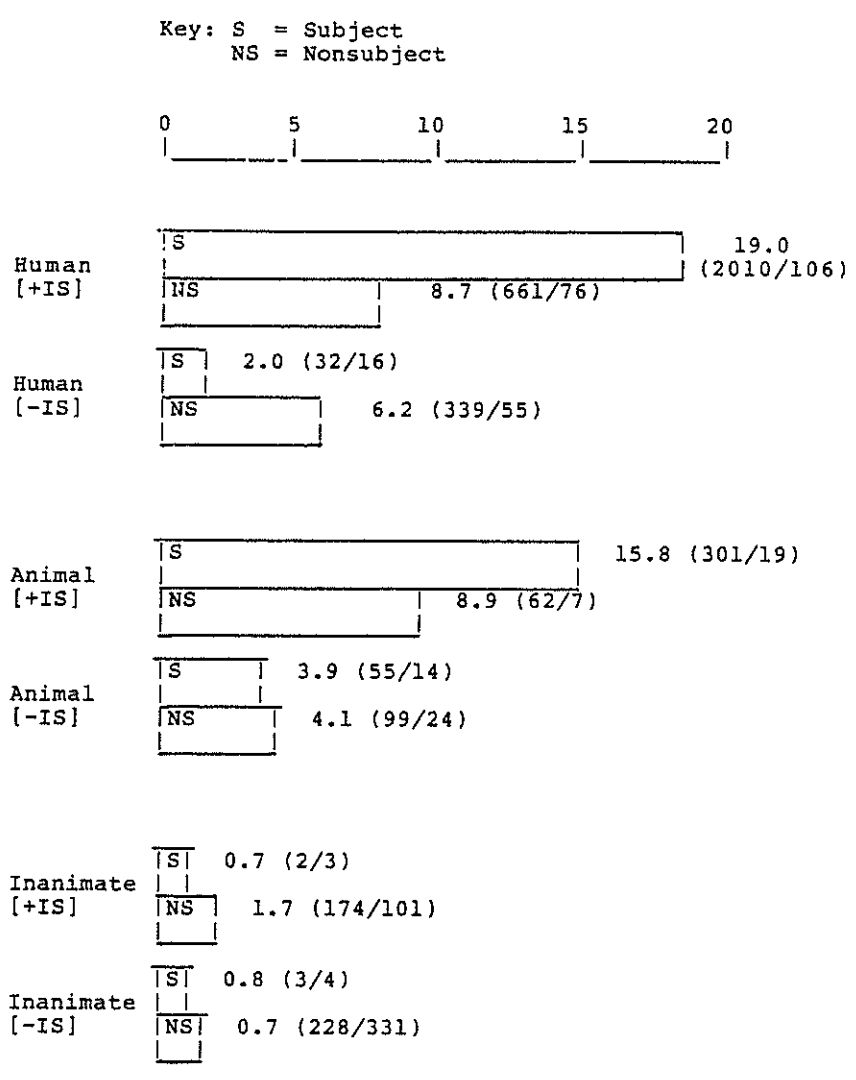


Figure 2.2. Average number of subsequent mentions of indefinites according to IS-marking, semantic category and syntactic status.[15]

With the occasional exception, to be discussed in full, the statistics in Figure 2.2 provide general confirmation of the predictions. Thus:

(17) Human indefinites, especially those bearing an IS on first mention, control the highest average number of subsequent discourse mentions.

(18) Indefinites introduced in the subject slot display a higher discourse life-span than their nonsubject counterparts.

The remainder of this chapter is now devoted to a detailed, category-by-category presentation of the supportive evidence summarized in Figures 2.1 and 2.2.

4. First-mention human indefinites

Consistent with the hypothesis, Figures 2.1 and 2.2 demonstrate, respectively, that a noticeably higher proportion of human arguments are marked with an Indefinite Specifier on initial mention, regardless of grammatical status, and that [+IS] human indefinites -- in this instance together with their animal counterparts, and in stark contrast to inanimates -- are invested with a relatively high degree of discourse-deployability. Let us now consider the findings in detail.

4.1. Subject human indefinites [+IS]

The overwhelming majority of human arguments occupying the subject role on first appearance are IS-marked -- 106 out of a total 122 counted (86.9%, cf. Figure 2.1) -- compared with a percentage of 57.6% for subject animal indefinites, and 42.9% for the restricted 3-member class of subject inanimates [+IS]. Just as significantly, Figure 2.2 shows that human indefinites enjoy the longest life-span of all, averaging a hefty 19.0 additional mentions before disappearing, although this value is not significantly higher than the mean of 15.8 computed for subject animal indefinites (cf. section 5.1 for a suggested explanation). Subject human indefinites thus appear to represent a category of highly deployable, highly codable, ongoing discourse "topic", with special communicative value. Once introduced in typically presentative constructions, they soon move to transitive environments, taking control of the verbalized actions and events which constitute the real stuff of the text.[16] The observed correlations are not surprising. Numerous studies have independently shown that humans generally find their fellow humans to be more intrinsically and perceptually salient/noteworthy (cf., inter alia, Chafe 1976, 1980a; Givón 1976, 1983; Li and Thompson 1976; Duranti and Ochs 1979; Kirsner 1979:360ff.; Du Bois 1980b:228ff.); and claims

regarding the elevated status of "subject" (however this notion is defined for a given language), e.g. its correlation with such properties as intentional agency, definiteness, humanness, discourse salience etc., have a wealth of supportive documentation (cf., inter alia, García 1975; Keenan 1976; Li and Thompson 1976; Fillmore 1977; Kirsner 1979; Zubin 1979; Bernardo 1980; and Givón 1983). These two variables -- humanness and subjecthood -- translate, within the context of integrated discourse, into a high degree of deployability which in turn correlates with exploitation of the Indefinite Specifier. Example (19) is typical:

(19) a. **wata** rana **wadansu** 'ya-m birni
IS day IS people-of city-SUBJ
ø-na zaune,
CONC-ø-IMPFV sit-STATIVE
b. sai ga **wani** bakauye ø
then PRESENT IS villager-SUBJ STAT-ø
tafe...
travel-STAT

(Imam 1971:36)

'One day some city folk were sitting around when along came a (certain) villager...'

The IS-bearing subject human indefinites wadansu 'yam birni 'some city folk' in (19a) and wani bakauye 'a (certain) villager' in (19b) are both highly salient discourse arguments, deployed as active participants until the end of the story, and IS-marking of such indefinites is the nearest thing to an absolute "rule" emerging from the

corpus under consideration. Fragment (20) illustrates the stylistic presentative construction mentioned in fn. 2 above, a literal gloss of which would be "one (impersonal) did X":

(20) A ciki-n zamani-n da
At inside-of period-of before-ADV
an yi wani sarki a nan gabas
IMPERS-PFV do IS emir-DO in there east-ADV
wai shi Shahruzzaman.
called he-IND PRO Shahruzzaman

(Imam 1970:167)

'Once upon a time there was a (certain) emir in the east called Shahruzzaman.'

This construction is an extremely common means of introducing highly deployable referents at the beginning of a story -- both singular and plural, and especially salient humans like sarki 'emir' in (20). Since, moreover, it involves the vacuous use of the perfective impersonal auxiliary an with the verb yi 'do', and is paraphreaseable with an existential-presentative introduced by the akwai 'there is/are' predicator, I have chosen to treat presentative NP's in such frames as subject arguments, i.e. despite the fact that they occupy the direct object slot formally speaking.

Finally, it is worth noting that the Indefinite Specifier can occur in isolation as an indefinite pronoun -- subject and nonsubject -- in which case it shares with its English counterpart 'someone' the property of introducing

a relatively unimportant referent, i.e. one not destined to play any key, protracted role in the discourse (Du Bois 1980b:221).[17] Excerpt (21) illustrates this usage:

(21) a. ...har duk sa'ad da aka fadi
so that every time REL IMPERS-PFV say
suna-n-sa
name-of-his-DO
b. sai ka ji
then you-SUBJCTV hear
c. wani ya amsa da
IS-SUBJ he-SUBJNCTV answer with
kirari...
praise-epithet-ASSOC

(Imam 1970:10)

'...so that whenever his name was mentioned you would hear someone respond with a praise-epithet...'

The precise identity of the individual referred to by the isolate subject IS wani 'someone' in (21c) is unknown to the encoder and/or unimportant to the communication, and no further mentions occur.[18]

4.2. Subject human indefinites [-IS]

This class is certainly in a minority -- 16 (13.1%) tokens recorded out of a total of 122 (cf. Figure 2.1). Figure 2.2, moreover, reveals a low average deployability -- 2.0 subsequent mentions, i.e. significantly less than the 19.0 calculated for the IS-marked variants, and lower even than the average 3.9 scored by the corresponding subject animal [-IS] subclass. This insubstantial average

-- it it the smallest for any human category -- is due to the fact that all but two of the 16 type-tokens were bare nouns stems, either nonspecific generics of the kind exemplified in (4b), or categorizing predicate nominals such as those noted in (5d, 6a, 7c), i.e. indefinites with a reduced potential for autonomous deployability. Fragments (22-23) illustrate the two cases which are not exemplars of these classes:

(22) a. 'Ya ɗaya kaɗai gare shi
daughter one-SUBJ only to him-DO PRO
b. an ko yi mata aure.
IMPERS-PFV and do to her-IO PRO marriage

(Imam 1971:1)

'He had only one daughter (lit. 'One daughter only to him'), even she was married...'

Example (22) is the continuation of fragment (10) above, in which the author is describing the emir's sorrow at not having any close kin. It is important to note that although the presentative indefinite subject NP 'ya ɗaya 'one daughter' in (22a) bears no IS on first mention, it is modified with the postposed numeral ɗaya 'one'.[19] We encounter a similar case in fragment (23). The addressee has been asked to list the items which were stolen from his pocket and answers facetiously:

(23) ...da kyanwa biyar, da wani ɗa-n
and cat five-SUBJ and IS DIMIN-of
jirgi-n ƙasa,
boat-of ground-SUBJ

da wata makaranta ta elementare a ciki,
and IS school of elementary-SUBJ at inside-ADV
da yara uku...
and children three-SUBJ

(Imam 1970:141)

'...and (there were) five cats, and a train, and an elementary school inside, and three children...'

Again we have an indefinite subject human NP -- yara uku 'three boys' at the end of the list in (23a) -- which is not IS-marked but which does carry a modifying numeral uku 'three'.[20] The generalization seems to be, therefore, that human indefinites, especially subjects, differ from their nonhuman (animal and inanimate) counterparts in that they are generally too salient to allow first mention with a simple nonmodified stem, with the concomitant possibility, as noted in 2.1, of a nonspecific generic interpretation. Their prominence thus favors individual categorization of some kind. For subject arguments, as noted in 4.1, this is typically achieved by the highly conspicuous strategy of IS-marking; and if this option is not chosen, then human subjects (excluding the generic and predicate nominal cases) are modified with some other qualifier, e.g. a numeral (22-23), or adjectival-nominal (24). When we come to consider the class of nonsubject human indefinites [-IS] in section 4.4 below, we shall note an even greater proportion of tokens from this [-IS] but [+modifier] category.

Finally, in this regard, I found one other example in Imam's (1966) final volume:

(24) A **wani** gari wai shi Samana
In IS town-ADV called it-IND PRO Samana
an yi baƙo-n farke-DO wanda
IMPERS-PFV do stranger-of trader REL
Ø-ke da sukuni ƙwarai da gaske...
CONC-Ø-IMPFV with leisure-ASSOC really-ADV

(Imam 1966:205)

'In a town called Samana there was a (certain) immigrant long-distance trader who was a man of real leisure...'

Again, the [-IS] indefinite human subject nominal farke 'trader' in (24) is modified, and so partially specified, both by a preposed adjectival nominal baƙo-n 'stranger of', and a following relative clause beginning with the masculine singular relative pronoun wanda 'who'. It is important to note too that we cannot relate the absence of an IS in such cases to a low degree of deployability, since the indefinite referent in (24a) for example, is manipulated as a highly salient and pivotal character until the very end of the story.

4.3. Nonsubject human indefinites [+IS]

We note in Figure 2.1 that although this category is in the majority -- 58.0% (76 out of 131), compared with 42.0% [-IS] -- this figure contrasts with the 86.9% computed for the corresponding subject human indefinites

[+IS]. Observe, however, that this 58.0% figure is higher than the 22.6% and 23.4% recorded for the corresponding animal and inanimate categories respectively. Figure 2.2 also demonstrates that nonsubject human indefinites display a substantially lower persistence rate than their subject counterparts -- 8.7 subsequent mentions vs. 19.0 respectively, with the 8.7 value comparable to the 8.9 figure for IS-marked nonsubject animals, both tallies being far greater than the 1.7 average recorded for the IS-marked nonsubject inanimate category. Notice too that the figure of 8.7 subsequent mentions is also marginally higher than the 6.2 calculated for the 55 [-IS] tokens in this class (cf. section 4.4). Fragments (25) and (26) are typical examples of the nonsubject human [+IS] subclass:

(25) a. Da Ø ya-na da uku,
Previously-ADV SUBJ-Ø he-IMPFV with three-ASSOC
b. sai Ø ya auro **wata** hatsabibiya-r
then SUBJ-Ø he-PFV marry IS sorceress-of
karuwa...
prostitute-DO

(Imam 1971:58)

'Previously he had three (wives), then he married a (certain) sorceress prostitute...'

(26) a. ...Ø ya-na kewaya lambu
SUBJ-Ø he-IMPFV go round garden-DO
b. sai Ø ya tarad da **wani** tsoho,
then SUBJ-Ø he-PFV come upon IS old man-ASSOC
c. ashe shi ne mai lambu-n.
in fact he-IND PRO SUBJ COP one with garden-DET

(Imam 1970:177)

'...he was going round the garden when he came upon a (certain) old man, in fact he was the owner of the garden.'

The two IS-coded nonsubject human indefinites -- the direct object wata hatsabibiyar karuwa 'a (certain) sorceress prostitute' in (25b) and the associative object wani tsoho 'a (certain) old man' in (26b) -- are both deployed as autonomous participants, the former until the end of the story and the latter appearing on 7 more occasions including (26c).[21]

Fragments (27) and (28) instantiate an additional noteworthy function of the IS:

(27) a. ...sai ∅ ya gaya wa
then SUBJ-∅ he-PFV tell to
wata ƙanwa-r uwa-r-sa
IS younger sister-of mother-of-his-IO
b. ∅ ta tafi
SUBJ-∅ she-SUBJNCTV go
c. ∅ ta kira liman.
SUBJ-∅ she-SUBJNCTV call imam-DO

(Imam 1970:128)

'...then he told a (certain) younger sister of his mother to go and call the imam.'

The IS-marked indefinite NP wata ƙanwar uwarsa 'a (certain) younger sister of his mother' in (27a) is of some interest since the selection of the IS wata is not, in this particular case, motivated by considerations of participant-deployability, but by the semantic-pragmatic necessity of maintaining a distinction between definite

and indefinite reference. In possessive constructions of the kind instantiated here, overt IS-coding of the possessee argument -- NP_1 in Hausa -- is obligatory if an indefinite reading is desired, for the absence of an IS in this environment would force a definite interpretation, i.e. 'the younger sister of his mother'. Compare, in this regard, English 'his friend [+definite]' vs. 'a friend of his [-definite]'.

Extract (28) -- this time from one of the oral narratives -- illustrates this same principle:

(28) Wata rana ne dai muka je biki-n
IS day COP well we-PFV go party-of
wani aboki-n-mu...
IS friend-of-us

(Speaker 1-2)

'Well one day we went to a party of a (certain) friend of ours...'

4.4. Nonsubject human indefinites [-IS]

In contrast to their subject counterparts, and in keeping with the predictions, a fair proportion of nonsubject human indefinites -- 55 out of 131 (42.0%) -- carried no IS on first mention, with the corresponding percentages for nonhumans even higher -- 77.4% and 76.6% respectively for animals and inanimates (Figure 2.1). In addition, the class of [-IS] nonsubject humans exhibits a decay rate,

i.e. disappearance from the text, which is marginally faster than the [+IS] forms -- 6.2 vs. 8.7 subsequent mentions (Figure 2.2) -- both of which values are higher than those computed for the nonsubject nonhuman [-IS] categories which again manifest the lowest discourse-deployment averages -- 4.1 and 0.7 subsequent mentions for animals and inanimates respectively.

Recall now the proposal, outlined and and briefly exemplified in section 4.2, that the IS-marking requirement for human indefinites may be relaxed if the NP in question has the following property: it is already specified by an accompanying modifier of some kind. When one scrutinizes the available data it turns out that this marking feature is present on almost half of the attested 55 nonsubject human [-IS] tokens. The rest were cases of bare, nonmodified stems from the following categories: generics, again as noted above; negative (associative) predicate indefinites, as in (10-11); and relational kin-terms. I found only one token not selected from one (or more) of the above categories:

(29) a. Don wata rana wani d̂a-m fashi
because IS day IS DIMIN-of robbery-SUBJ
ya tsare farke,
he-PFV intercept trader-DO
b. ø ya kama farke-n
SUBJ-ø he-PFV seize trader-DET-DO
c. ø ya kayar ø...
SUBJ-ø he-PFV knock down ASSOC-ø

(Imam 1970:129)

'Because one day a (certain) highway robber intercepted a long-distance trader, seized the trader and knocked (him) down...'

Clause (29a) begins with obligatory IS-marking of the orientational time-phrase wata rana 'one day', followed by the (near) obligatory IS-marking of the first-mention subject human indefinite wani d̂am fashi 'a (certain) highway robber', and then a nonsubject human indefinite farke 'long-distance trader' which carries neither an IS nor an accessory modifier of any kind. Again, we cannot appeal to the notion of non-deployability to explain this case, for the referent in question in fact controls several subsequent mentions, including full nominal with a determiner (cf. Chapter 4) in (29b), and zero anaphoric reference (cf. Chapter 3) in (29c). I would suggest, instead, that it is the stylistic desire to avoid the clustering of too many IS-coded NP's which perhaps explains the nonoccurrence of an IS here. As observed by Skinner (1974:253), major concentrations of IS-marked NP's tend to occur either at the very beginning of the narrative itself or, as in (29a), at story-internal episode boundaries. These are obviously points at which new and salient characters are introduced, and where important spatiotemporal settings are specified -- cf. fragment (40) below where we encounter a fairly dense clustering of IS's. However, IS-marking of every first-mention indefinite in such environ-

ments would, I believe, act to overwork this morpheme and so dilute the force of its distinctive particularizing/individuating semantics.

Turning now to the more frequently-encountered referential form -- nonsubject indefinites which carry no IS but which do bear some other type of modifier -- excerpts (30-34) are illustrative:

(30) ...ø ya ƙara masa
SUBJ-ø he-PFV increase to him-IO PRO
da bayi bakwai...
with slaves seven-ASSOC

(Imam 1971:9)

'...he added seven slaves for him...'

(31) a. Abi-n-ka da mai sule,
thing-of-you with one with shilling-ASSOC
b. kafin kwana bakwai ø ya nemi
before day seven SUBJ-ø he-PFV look for
'ya-m mata uku 'ya-n gaske,
DIMIN-of women three-DO DIMIN-of truth
c. aka ba shi ø.
IMPERS-PFV give him-DO PRO DO-ø

(Imam 1971:62)

'You know what a rich man is like, within seven days he had looked for three real young girls, and he was given (them).'

The two human nonsubject indefinites we are concerned with are the associative object NP bayi bakwai 'seven slaves' in (30) and the direct object NP 'yam mata uku 'yan gaske 'three real young girls' in (31b). Neither carries any IS, but they are both modified with a postposed numeral --

bakwai 'seven' and uku 'three' respectively -- and in (31b) there is additional specification with the appositional phrase 'yan gaske 'real/true'. My basic contention is, therefore, that the presence of such modifiers acts to partially particularize/individuate/specify etc. the indefinite so-marked, and sufficiently to allow nonuse of an additional specifier like the IS.[22]

Examples (32-34) illustrate the use of modifiers other than numerals:

(32) a. ...sai kaka-n ya tafi
then grandfather-DET-SUBJ he-PFV go
b. ø ya gaya wa sarki,
SUBJ-ø he-SUBJCTV tell to emir-IO
c. ø ya ce
SUBJ-ø he-PFV say
d. ø ya-na da jika wanda ba
SUBJ-ø he-IMPFV with grandson REL NEG EXIST
abi-n da za a yi masa
thing-DET REL FUT IMPERS do to him-IO PRO
e. ø ya kaɗu.
SUBJ-ø he-SUBJNCTV take fright

(Imam 1970:21)

'...then the grandfather went to tell the emir, and said he had a grandson who would never take fright whatever was done to him.'

In (32d), the indefinite associative NP jika 'a grandson' -- the hero of the story -- is modified by a following relative clause beginning with the masculine singular relative pronoun wanda. Similarly, in (33a):

(33) a. ...baya-n wata tara ø ta haifi ɗa
after-of month nine SUBJ-ø she-PFV bear son

namiji wanda kyau-n-sa ba shi da
male-DO REL beauty-of-his-SUBJ NEG it-IMPFV with
iyaka.
limit-ASSOC

(Imam 1971:2)

'...after nine months she bore a son whose beauty was beyond description.'

the direct object human indefinite ɗa 'a son' is qualified with a postposed (collocational) adjectival nominal namiji 'male' and a following relative clause again. Finally, in this regard, consider extract (34):

(34) a. ...sai Ø ya ga 'ya-r tsohuwa
then SUBJ-Ø he-PFV see DIMIN-of old woman-DO
da kaya-n tulu
with loads-of water pot-ASSOC
b. duk Ø ta jiƙa,
all SUBJ-Ø she-PFV be wet
c. sai makyarkyata Ø ta-ke yi...
only trembling SUBJ-Ø she-IMPFV do-VN

(Imam 1970:80)

'...then he saw a small old woman with a load of water-pots, and she was wet through and just trembling...'

where the direct object indefinite referent 'yar tsohuwa 'a small old woman' in (34a) is not marked with any IS because it is already specified by the preposed diminutive 'yar (lit. 'daughter of').

Notice that several of the foregoing nonsubject human indefinites [-IS] are kin-terms -- cf. (32) and (33). I think there is evidence to suggest that this particular class has features which perhaps distinguish it from other

lexical classes. There are cases in the corpus, for example, where indefinite kin-terms bear no modifier of any description on first mention, IS or otherwise. Consider, therefore, fragment (35):

(35) a. To, amma ko da ya ke sarki ya yi murna
OK but although emir-SUBJ he-PFV do joy
ƙwarai
really-ADV
b. da Ø ya sami jika,
when SUBJ-Ø he-PFV get grandson-DO
c. duk da haka murna-r-sa ragaggiya ce...
nonetheless joy-of-his-SUBJ reduced COP

(Imam 1971:1)

'OK, but although the emir was overjoyed that he had got a grandson, his joy was nonetheless incomplete...'

where the direct object indefinite jika 'a grandson' in (35b) bears no modifier, IS or otherwise, despite the fact that the referent is destined to play a key part in the unfolding plot. I think the answer to this problem may have something to do with the fact that such inalienable kin-terms usually follow upon prior mention of the associated possessor argument. Although first mention of an indefinite kin-term remains nonidentifiable to the decoder, overt mention of the possessor activates an appropriate human-based frame allowing specification of the possessed kin-referent which is partial but nonetheless sufficient to allow use of the bare stem. The referent may be said to "belong" in the context of a whole-part rela-

tionship, and so the pressure to IS-mark is relaxed for this class of items.[23] Possible independent support for the proposal that possessee kin-terms are in a sense intrinsically categorized/individuated via their intimate and consistent association with a possessor argument stems from the language-specific fact that they may even bear a possessive pronoun suffix on first mention, e.g.,

(37) a. Akwai **wani** mutum wai shi Kalala,
EXIST IS man-SUBJ called he-IND PRO Kalala
b. Ø ya-na da mata-r-sa...
SUBJ-Ø he-IMPFV with wife-of-his-ASSOC

(Imam 1970:15)

'There was a man called Kalala, he had his (a) wife...'

(38) a. Domin wai akwai **wani** masassaki
because allegedly EXIST IS carpenter-SUBJ
a Katako,
in Katako-ADV
b. Ø ya-na da 'ya-r-sa mai
SUBJ-Ø he-IMPFV with daughter-of-his-ASSOC with
shekara biyar da haihuwa, suna-n-ta Halima.
year five with birth name-of-her Halima

(Imam 1971:94)

'For it is said there was a (certain) carpenter in Katako, he had his (a) daughter five years of age, her name Halima.'

The associative-possessee kin-referents in question -- matarsa 'his wife' in (37b) and 'yarsa 'his daughter' in (38b) -- both carry a masculine possessor pronoun suffix -sa, despite the fact that they are first-mention indefinites.[24]

If the above proposals have any merit to them, then they may be related to the fact that, in English, reference to such entities as body-parts, clothing, sub-locales etc., allow partial categorization via "frame" membership (Du Bois 1980b), permitting initial-mention coding with a definite article or possessive pronoun.

5. First-mention animal indefinites

In terms of IS-support and discourse-deployability, animal indefinites seem to occupy a position roughly intermediate between their highly salient, highly codable human counterparts and the less salient, low codable inanimate class. Thus, Figure 2.1 shows that 19/33 (57.6%) of animals introduced in the subject role are IS-marked, compared with 86.9% recorded for subject human indefinites and 42.9% for inanimates. Notice too that, as was the case with human indefinites, and again according to the predictions, the nonsubject animal category is less frequently IS-marked -- 22.6% of the token count. And turning to the statistics in Figure 2.2, we note again a systematic interpredictability between the discourse-deployability of first-mention indefinites, their syntactic status, and the strategy of IS-marking. Let us now consider the data, category by category.

5.1. Subject animal indefinites [+IS]

As already noted, this category has the highest percentage of IS-marked tokens -- 57.6%. A glance at Figure 2.2 reveals an additional noteworthy fact: IS-marked subject animal indefinites have a persistence rate not far short of that computed for subject humans -- means of 15.8 vs. 19.0 subsequent mentions respectively. Closer inspection of the 19 type-tokens indicates, however, that this relatively high average has a natural explanation which ties in nicely with earlier claims regarding the special discourse status of subject arguments, especially human subjects. Thus, 8 of the 19 attested cases involved personified animals occurring as the major actors in animal-based fables, the 'mouse' and the 'crownbirds' in (39) and (40) below being typical examples. These animals take on quasi-human traits, including the power of speech in many cases, and the active-agential ability to consciously control events and actions within the story, and so attract the morphological trappings usually associated with human arguments, especially subjects, i.e. IS-coding on first mention.[25] Fragments (39) and (40) provide exemplification:

(39) a. A gindi-n itaciya-n nan akwai **wani** ɓera,
at foot-of tree-DET DEM-ADV EXIST IS mouse-SUBJ
b. a ciki-n kogo-n-ta kuma ashe akwai
at inside-of hollow-of-it-LOC and in fact EXIST

wani muzuru...
IS tom-cat-SUBJ

(Imam 1971:47)

'At the foot of that tree there was a (certain) mouse, and in its hollow there was a (certain) tom-cat..'

(40) a. Wata rana **wadansu** gauraki guda biyu
IS day-ADV IS crownbirds unit two-SUBJ
suka gane **wani** itace-n ɓaure a baki-n
they-PFV notice IS tree-of fig-DO at edge-of
wani rami.
IS stream-LOC

(Imam 1971:76)

'One day two crownbirds noticed a (certain) fig-tree at the edge of a (certain) stream.'

Both the subject animal indefinites -- wani ɓera 'a (certain) mouse' in (39a) and wadansu gauraki guda biyu 'two crownbirds' in (40a) -- are provided with an IS and persist as major referents in their respective texts.

5.2. Subject animal indefinites [-IS]

Figure 2.1 shows that 42.4% (14/33) of subject animal indefinites were not IS-marked, i.e. a substantially higher percentage than the 13.1% (16/122) calculated for the corresponding human category, but still lower than the 57.1% recorded for inanimates. And consistent with the hypothesis, Figure 2.2 reveals that subject animal indefinites [-IS] decay at a much faster rate than their [+IS] counterparts -- 3.9 vs. 15.8 subsequent discourse men-

tions -- though in this case the average is in fact fractionally lower than the 4.1 computed for the nonsubject [-IS] category. Extracts (41) and (42) illustrate:

(41) a. Wata rana yautai ya-na kiwo
IS day-ADV nightjar-SUBJ it-IMPFV feed
a baya-n gari
at back-of town-LOC
b. sai tarko ya kama shi.
then trap-SUBJ it-PFV catch him-DO PRO
c. ø Ya yi
SUBJ-ø it-PFV do
d. ø ya yi
SUBJ-ø it-PFV do
e. ø ya kuɓuce,
SUBJ-ø it-SUBJNCTV escape
f. ø ya kasa.
SUBJ-ø it-PFV fail

(Imam 1971:109)

'One day a nightjar was feeding behind a town
when a trap caught it. It tried and tried
to escape but failed.'

(42) a. ø Ya duba sama
SUBJ-ø it-PFV look upwards-ADV
b. don ø ya hau,
in order to SUBJ-ø it-SUBJNCTV climb up
c. sai ø ya ga
then SUBJ-ø it-PFV see
d. shirwa ø-na kewaya,
black kite-SUBJ CONC-ø-IMPFV circle-VN
e. ø ta-na jira
SUBJ-ø she-IMPFV wait-VN
f. ø ya motsa
SUBJ-ø it-SUBJNCTV move
g. ø ta fyauce shi.
SUBJ-ø she-SUBJNCTV pounce on it-DO PRO

(Imam 1971:47-48)

'It (the mouse) looked up so that it could climb up,
when it spotted a black kite circling around,
waiting for it to move and she would pounce on it.'

Neither of the two animal indefinites in the above examples -- yautai 'a nightjar (bird)' in (41a) and shirwa 'a black kite' in (42d) -- is assigned an IS, despite the fact that they are both introduced as clausal subjects and both are extensively deployed in the following text, with the 'nightjar' in fact assuming human-like traits. It seems to be the case, therefore, that unlike the human indefinite [-IS] examples documented in 5.2 and 5.4, where I suggested that such arguments were generally too salient to allow coding via a simple nominal, animal indefinites occur as bare nonmodified stems with greater frequency. If we can assume a salience hierarchy, then animal referents are located at a position somewhat lower than human referents, with inanimate entities plotted at the lowest extremes of the scale, and all these relative positions have consequences for morphological coding. Had the subject animal indefinites in (41a) and (42d) been human, however, and assuming that specific (non-generic) reference was intended, then marking with an IS or some other modifier would have been compulsory. Such examples are important because they provide support for the claim that it is the combined effect of syntactic status and semantic class which conditions IS-marking of indefinites -- human indefinites in general attract the IS with far greater frequency because they are perceived as more salient, noteworthy entities.

5.3. Nonsubject animal indefinites [+IS]

Figure 2.1 shows that only 22.6% (7/31) of nonsubject animal indefinites are IS-marked, compared with 58.0% for nonsubject humans, and much closer to the 23.4% recorded for nonsubject inanimates [+IS]. Recall too that a substantially higher proportion of subject animals are IS-coded on first mention -- 57.6%. In addition, Figure 2.2 reveals once again that the IS-marked nonsubject animal class has a weaker persistence rate than the IS-marked subject forms -- 8.9 vs. 15.8 subsequent mentions -- but a stronger rate of survival than the nonsubject [-IS] class (4.1) and the corresponding nonsubject inanimate [+IS] category (1.7). Fragments (43) and (44) are typical:

(43) a. ø Ya-na kuma da **wani** aku...
SUBJ-ø he-IMPFV and with IS parrot-ASSOC
b. ...sai ø ta kama aku,
then SUBJ-ø she-PFV seize parrot-DO
c. ø ta fige shi da rai...
SUBJ-ø she-PFV pluck him-DO PRO with life-ADV

(Imam 1971:41)

'And he (the merchant) had a (certain) parrot...then she (the merchant's wife) seized the parrot and plucked him alive...'

(44) a. ...sai ø ta iske
then SUBJ-ø she-PFV find
b. tarko ya kama **wani** namiji-n gwara,
trap-SUBJ he-PFV catch IS male-of sparrow-DO
c. ø ta ga
SUBJ-ø she-PFV see

d. mata-r ta duƙa
mate-DET-SUBJ she-PFV bend down
e. ø ta-na tsuntsunke tarko-n
SUBJ-ø she-IMPFV peck at trap-DET-DO
da baki-n-ta
with beak-of-her-ASSOC
f. miji-n ya fita.
male-DET-SUBJ he-SUBJNCTV escape

(Imam 1970:169)

'...then she (the girl) found that a trap had ensnared a (certain) male sparrow, and she saw the mate bend down pecking at the trap so that the male could escape.'

The IS-marked associative-possessed indefinite **wani aku** 'a (certain) parrot' in (43a) continues in personified form until the end of the story (it doesn't die despite the attentions of the wife!), and the similarly IS-marked direct object animal indefinite **wani namijin gwara** 'a (certain) male sparrow' in (44b) controls 5 subsequent mentions, including the subject reference in (44f).

5.4. Nonsubject animal indefinites [-IS]

This subcategory has the highest percentage of [-IS] forms -- 77.4% (cf. Figure 2.1) -- a value comparable to the equally low-codable class of nonsubject animal indefinites (76.6%), and higher than the 42.0% recorded for the more salient nonsubject human cases. And Figure 2.2 shows that together with their subject counterparts discussed in 5.3 above, the class of nonsubject animal indefinites [-IS] has a lower persistence rate than the corresponding

IS-marked nonsubjects -- 4.1 vs. 8.9 subsequent mentions, again somewhat intermediate between the human (6.2) and inanimate (0.7) categories. Fragments (45) and (46) are typical:

(45) a. ...sai ga wani Balarabe
then PRESENT IS Yoruba-SUBJ
Ø ɗauke da aku ciki-n
STAT-Ø carry-STAT with parrot-ASSOC inside-of
keji.
cage-LOC
b. Ko da Musa ya ɗaga ido
As soon as Musa-SUBJ he-PFV raise eye-DO
c. Ø ya gan shi...
SUBJ-Ø he-PFV see him-DO PRO

(Imam 1971:5)

'...then there appeared a (certain) Yoruba man, carrying a parrot in a cage. As soon as Musa raised his eyes he saw him (the parrot) ...'

In (45a) the associative indefinite bare nominal aku 'a parrot' carries no specifier of any description, despite the fact that the referent is destined to play an important role in the unfolding narrative. As I suggested in section 5.2, moreover, this kind of "deindividuating" coding would have been most unnatural, not to say unacceptable, had the referent been a member of the highly prominent human class. Compare this example too with the IS-marked associative indefinite wani aku 'a (certain) parrot noted in (43a). Similarly in (46):

(46) Wani mutum Ø-na da alfadari
IS man-SUBJ CONC-Ø-IMPFV with mule
da jaki.
and donkey-ASSOC

(Imam 1971:22)

'A (certain) man had a mule and a donkey.'

neither of the associative-possessed animal indefinites alfadari 'a mule' and jaki 'a donkey' bears an IS or other modifier, though both arguments persist as autonomous referents into the story.

6. First-mention inanimate indefinites

The prevailing pattern observed for inanimate referents is that they serve to provide background information about human participants and the activities they are engaged in. In the general case, inanimates are not sufficiently salient to be deployed as autonomous, independent discourse arguments, at least to the same extent that humans, and to a lesser degree animals, are manipulated, a feature also noted for English and Sacapultec (Du Bois 1980b, 1981). Inspection of Figure 2.1 shows, therefore, that only a handful of subject inanimate indefinites was in fact recorded -- 7 tokens in all, 3 [+IS] and 4 [-IS] -- and that the vast majority -- 432 tokens -- occurred as nonsubject indefinites.[26] And Figure 2.2 furnishes additional evidence of this relative lack of salience -- inanimate inde-

finites, whatever their grammatical status on first mention, and regardless of whether they were assigned an IS -- display by far the lowest averages for discourse persistence.

6.1. Subject inanimate indefinites [+IS]

A mere three tokens of this relatively rare subclass were recorded in the corpus investigated, now exemplified in (47) and (48) (the third token occurs in (54a) below):

(47) a. ø Ya-na zuwa
SUBJ-ø he-IMPFV come-VN
b. sai ø ya tarad da
then SUBJ-ø he-PFV find
c. **wani** jirgi ya zo,
IS boat-SUBJ it-PFV come
d. a-na ta fid da kaya daga
IMPERS-IMPFV continue take out loads from
ciki-n-sa.
inside-of-it-LOC

(Imam 1971:26)

'He was coming along and then found that a (certain) boat had arrived, and loads were being taken out of it.'

(48) Sai **wata** dabara ta faɗo masa...
then IS idea-SUBJ it-PFV fall to him-IO PRO

(Imam 1970:139)

'Then a (certain) idea came to him...'

Imam's intent seems to be to particularize/individuate the two subject inanimate indefinites in question -- wani jirgi 'a (certain) boat' in (47c) and wata dabara 'a

(certain) idea' in (48), though neither referent turns out to be especially salient, hence the low average of 0.7 subsequent mentions for this subclass (Figure 2.2). The first inanimate controls two further mentions, including the masculine singular possessive pronoun suffix /-sa/ in (47d), and the latter referent disappears immediately, though it does provide background motivation for the events of the following episode.

6.2. Subject inanimate indefinites [-IS]

This subcategory is as rare as the foregoing subject [+IS] class, and has a survival rate which is equally short -- a mean of 0.8 mentions. Clauses (41b) and (44b) both contain examples -- the indefinite inanimate tarko 'a trap' is clausal subject in both cases -- and extract (49) illustrates one of the few remaining tokens:

(49) a. Kafin ø ya kai
before SUBJ-ø he-SUBJNCTV reach
b. hadari-n ruwa ya taso,
storm-of rain-SUBJ it-PFV arrive
c. ruwa ya yi ta duka-n-sa...
rain-SUBJ it-PFV do continue beat-of-him-VN

(Imam 1970:80)

'Before he reached (the emir's palace) a rainstorm came and the rain beat down on him continuously...'

None of the documented subject inanimates [-IS] -- hadarin ruwa 'a rainstorm' in (49b) or the examples in (41b) and

(44b) -- enjoy appreciable life-spans.

6.3. Nonsubject inanimate indefinites [+IS]

Figure 2.1 shows that only 23.4% of nonsubject inanimate indefinites (101/432) were IS-marked, a figure which is comparable to the 22.6% calculated for the corresponding animal category, but less than half the 58.0% recorded for the more salient nonsubject human indefinites [+IS]. Figure 2.2, moreover, indicates an average rightwards persistence of 1.7 clauses -- higher than the 0.7 calculated for the [-IS] equivalents, but significantly lower than the averages computed for the corresponding human (8.7) and animal (8.9) nonsubject indefinites [+IS]. Consider examples (50) and (51):

(50) a. Aka kawo **wani** zobe na zinariya
IMPERS-PFV bring IS ring of gold-DO
b. aka sa masa Ø a hannu...
IMPERS-PFV put to him-IO PRO DO-Ø on hand-ADV

(Imam 1970:90)

'They (someone) brought a (certain) gold ring and placed (it) on his hand...'

(51) a. Jim kaƊan Ø ya dawo da **waƊansu**
later little SUBJ-Ø he-PFV return with IS
ganyaye,
leaves-ASSOC
b. Ø ya ba sarki-n Sirika Ø,
SUBJ-Ø he-PFV give emir-of Sirika-DO DO-Ø
c. Ø ya ce
SUBJ-Ø he-PFV say

d. a dafa su...
IMPERS-SUBJNCTV boil them-DO PRO

(Imam 1970:57)

'A little later he came back with certain (kinds of) leaves, gave (them) to the emir of Sirika and said that they should be boiled...'

The direct object indefinite wani zobe na zinariya 'a (certain) gold ring' in (50a) and the associative indefinite waƊansu ganyaye 'certain (kinds of) leaves' in (51a) are both particularized with IS's, and both control a number of additional mentions -- the 'ring' makes 17 further appearances, and the 'leaves' are mentioned on two further occasions, as a zero-anaphoric direct object in (51b) and a direct object pronoun su 'them' in (51d) (cf. Chapter 3).

6.4. Nonsubject inanimate indefinites [-IS]

Use of a bare noun stem is by far the most common referential option for nonsubject inanimate indefinites -- 76.6% (331/432) of all such tokens occurring in this form (cf. Figure 2.1). This proportion is thus comparable to the 77.4% computed for nonsubject animals [-IS], with both figures somewhat higher than the 42.0% recorded for the more salient human category. And once again the prediction that the nonsubject [-IS] class will exhibit a generally more rapid rate of decay is borne out by the data in Figure 2.2 -- an average of 0.7 subsequent men-

tions, the lowest value for any indefinite subset. Non-subject inanimate and animal indefinites thus seem to display some common properties, in contrast, that is, to their more prominent human counterparts. Because they lack the relatively high degree of salience which generally characterizes human participants, there is consequently less pressure to surround them with morphology on first mention. Use of a bare, nonmodified stem is, therefore, a more acceptable and frequently-occurring referential option, and this deindividuating coding correlates in a regular fashion with relatively low degrees of discourse-deployability. (52) and (53) illustrate the phenomenon:

(52) a. ...Ø ya ɗauko gawa-r
SUBJ-Ø he-PFV pick up corpse-DET-DO
b. Ø ya kawo Ø gida,
SUBJ-Ø he-PFV bring DO-Ø home-LOC
c. Ø ya sami tsani
SUBJ-Ø he-PFV get ladder-DO
d. Ø ya gama Ø
SUBJ-Ø he-PFV join DO-Ø
e. Ø ya ɓoye Ø.
SUBJ-Ø he-PFV hide DO-Ø
f. Can da dare Ø ya tashi
then at night-ADV SUBJ-Ø he-PFV get up
g. Ø ya ɗauki gawa-n nan da
SUBJ-Ø he-PFV take corpse-DET DEM-DO and
tsani-n...
ladder-DET-DO

(Imam 1970:83)

'...he took the corpse and brought (it) home, and he got a ladder and put (them) together and hid (them). Later that night he got up and took that corpse and the ladder...'

(53) a. ...Ø ya-na jefa garma
SUBJ-Ø he-IMPFV throw hoe-DO
b. Ø ya-na cafe-wa Ø.
SUBJ-Ø he-IMPFV catch-VN DO-Ø

(Imam 1971:36)

'...he was tossing a hoe and catching (it).'

The two [-IS] direct object indefinites in question are tsani 'a ladder' in (52c) -- picked up on several more occasions, including two direct object zero references in (52d-e) and a direct object noun + determiner strategy in (52g) -- and garma 'a hoe' in (53a) -- referred to again with another direct object zero in (53b).[27]

7. The disjunctive 'other' reading of the IS

I would like to conclude this opening chapter with some brief remarks on an additional function of the IS briefly alluded to in section 1.1 -- its exploitation in specifying indefinite referents in contexts where English would use a disjunctive 'another X, other X's', e.g.,

(54) a. **Wata** hanya ta nufi gabas,
IS road-SUBJ it-PFV head east
b. **wata** ta nufi arewa,
IS-SUBJ it-PFV head north
c. **wata** ta nufi kudu...
IS-SUBJ it-PFV head south

(Imam 1970:36)

'One road headed east, another headed north, another headed south...'

The two anaphoric-pronominal occurrences of the IS -- feminine singular wata 'another (road)' in (54b-c) -- still serve to particularize their nonovert subject referents in exactly the same way that the same IS wata individuates the lexical NP indefinite in wata hanya 'one/a certain road' in (54a). When used in this fashion, the IS categorizes its referent and distinguishes it from the other members of the lexical class whose identity is denoted by the first IS-marked indefinite NP in the series. Perhaps, as I have suggested elsewhere (Jaggar 1983:417, fn.33), the principle which unifies these two functions of the IS relates to the notion of "participant-switching". More specifically, just as an IS-marked indefinite with an exclusive 'other' reading entails a switch from a previously mentioned member of the semantic class as in (54), so does IS-marking of a first-mention indefinite NP entail a switch, in this case from a zero linguistic environment where no preceding same-category entity has been expressed.[28] Seen in this light, the problem of accounting for the two apparently unrelated functions is possibly resolved.

Finally, it is of some interest to note that the IS can also be used in conjunction with the suffixal determiner (cf. Chapter 4) to modify an indefinite mention, e.g.,

(57) a. Wani ɓarawo ne ya shiga taska-r
IS thief-SUBJ COP he-PFV enter store-of

Kakaki...
Kakaki
b. ...sai wani ɓarawo-n kuma ya fado
then IS thief-DET-SUBJ and he-PFV burst in
ɗaki-n...
room-DET

(Imam 1971:19)

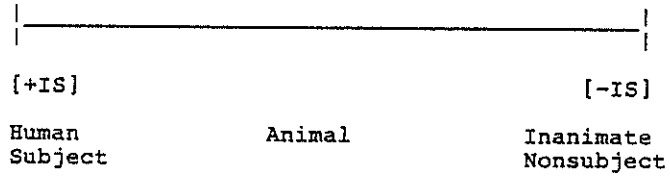
'A (certain) thief got into Kakaki's store...and then another thief burst into the room...'

In (57b) the IS wani is used to pick out and individualize an additional, but still nonidentifiable, token from the category established by the prior occurrence of the NP ɓarawo 'thief' in (57a). And the definite determiner /-n/ suffix on ɓarawon in (57b) serves, I would suggest, to modify the generalized and now identifiable category as a whole. This IS + NP + definite determiner conjunction is only possible, however, when the IS is being exploited to signal the disjunctive 'other X' reading.

8. Summary

This chapter has attempted to provide some answers to the fundamental linguistic question: "Why, in a given environment, is form X exploited, in preference to form Y?" In addressing this problem, we have seen a referential system not governed by hard-and-fast, absolute rules, but a system organized in terms of often strong, objectively quantifiable regularities in the packaging of first-mention

indefinites. The central claim was that the criterion underlying assignment of the Indefinite Specifier, complete with the suggested particularizing/individuating semantics, relates to the pragmatic notion of discourse-deployability or autonomy -- use of IS-marking covaries with higher degrees of discourse-deployability, and nonuse is a function of lower degrees of prominence. I demonstrated further that two dimensions had a determining influence upon the choice -- the lexicosemantic class and syntactic status of the first-mention indefinite -- and that together these factors render subject human participants the most salient and persistent of discourse arguments. Human indefinites are IS-marked with much greater frequency, and continue to be deployed well beyond their introduction clause, typically evolving into "global topics" within the discourse. The patterns allow the system to be characterized in terms of a graded continuum:



In the next chapter, we turn our attention to the domain of nonovert definite reference -- zero and pronominal anaphora -- and we shall note additional referential

choices which are conditioned by the same pragmatic factor of discourse salience.

Notes to Chapter 2

[1]Unless otherwise indicated, all Hausa citations in this dissertation are transcribed in accordance with Standard (Kano) Hausa orthography, and the glottalized consonants used are indicated with the "hooked letters" (lower case) ɓ, ɗ, ƙ, and (upper case) Ɓ, Ɗ, Ƙ. In addition, hyphens are used to indicate morpheme-boundaries, and items pertinent to the discussion in hand are in **boldface** type.

[2]For the purposes of this study, "subject" is a cover-term used, somewhat grossly, to denote the following items: (a) arguments controlling agreement on the preverbal tense-aspect auxiliary, where the lexical verb is often a highly presentative intransitive like 'sit', 'live', 'have', and 'come', e.g., in (1a), the IS-marked indefinite subject NP [+third person singular masculine] wani yaro 'a (certain) boy' governs agreement on the following auxiliary ya [+third person singular masculine, +perfective] which occurs before the finite verb zo 'come'; (b) predicate arguments occurring in nonverbal copular-equational constructions and in existential-presentative constructions introduced by the predicators ga or akwai; (c) indefinite arguments introduced via the stylistic mechanism An yi X (IMPERSONAL-SUBJ do-PFV X-DO) literally 'One did X', i.e. 'There once was an X...', where X is usually a first-mention human character, cf. example (20). And the label "nonsubject" encompasses the following postverbal syntactic categories: direct object; indirect object; prepositional object.

[3]With the notable exception of Hopper and Thompson (1983) who talk in terms of degrees of "manipulability" and the relation of this notion to the "high and low categoriality" of arguments, writers on the subject have generally preferred the term "referential" to characterize this discourse function. Du Bois (1980b:208-209), for example, defines as "referential" an NP which is used to speak about an object -- animate or inanimate -- as an object, with "continuous identity over time". And Givón (1982:86ff.) uses the label "(pragmatic-)referential" to designate arguments which have salience/importance in the ensuing discourse context. However, as pointed out by Hopper and Thompson (1983), Givón (1982:84) does make the important observation that the cross-language distribution of "reference coding properties" is not a function of the logic-based notion of "referentiality", i.e. the "existence" of a given entity in a real or imagined world of some kind; rather, it is linked to the "communicative intent of the speaker uttering the discourse, specifically...whether a particular individual argument (NP) is going to be important enough in the subsequent discourse, i.e. whether its specific identity is important, or only its generic type membership [original emphasis]". An additional drawback to the term "referential" relates to the fact that in Hausa, as is the case with many languages (Hopper and Thompson 1983), entities which are demonstrably "referential" in the logico-semantic sense that they "exist" in some universe, may nonetheless be coded as if they were not "referential", i.e. with reduced morphological coding, a feature which is explicable, in fact, in terms of the degree of discourse-deployability. It is in order to avoid these areas of potential confusion that I have elected to use the term "discourse-deployable".

[4]In this regard, one is immediately reminded of analogous facts reported for such diverse languages as Bemba and Modern Hebrew (Givón 1977:304, 1979b:99-100), Hungarian (Hetzron 1971:89-91), and Mandarin Chinese (Hopper and Thompson 1983), where the distribution of various language-specific "referential morphemes" is conditioned by whether an indefinite nominal denotes a specific entity which is discourse-manipulable. None of the above studies, however, are especially concerned with the impact of the properties of humanness and subjecthood on the use or nonuse of these formatives.

[5]The meaning of the various "ø" notations included in the morphological glosses will be explained in full when we come to consider zero anaphoric reference in the following chapter.

[6]For the benefit of those readers not acquainted with Hausa, the (masculine) copular element ne can be used either to modify an NP as in (4c), or to underline the "truth value" of the proposition underpinning a whole clause or sentence (Schachter 1966:39), as in (4a, g).

[7]Other spatiotemporal scene-setting expressions which collocate with the Indefinite Specifier include: wata rana 'one day' as in (13) below, wata shekara 'one year' (note the use of the particularizing numeral 'one' in the above English glosses), and wani lokaci 'sometimes'.

[8]It may be, therefore, that Hausa permits finely-graded "degrees of genericity", of the kind noted by Givón (1982:96ff.) for English and Spanish.

[9]Hopper and Thompson (1983:24) note an analogous feature in Mandarin, where the "classifier" used to mark manipulable indefinites may modify predicate nominals in similar contexts.

[10]As pointed out to me by John Du Bois (p.c.), this situation differs noticeably from English, where use of a plural indefinite marker "some" would be unacceptable, whatever the (motivational) importance of the predicate indefinite to the following text.

[11]I am aware that this type of "explanation" has involved recourse to qualitative justification of my claims, but I believe, with García (1983:191ff.), that such interpretive evidence can usefully complement quantitative validation in discourse analysis.

[12]Hausa thus differs from such languages as Israeli Hebrew, for example, where, according to Givón (1977:304, 1982:85ff.) the "referential-indefinite" morpheme is prevented from marking the object of a negative sentence.

[13]Although the IS has no etymological connection with the numeral 'one' (Newman 1972), use of this possible corresponding gloss does serve to underscore the individuating/particularizing function of the IS.

[14]The corresponding tallies from the four spontaneously-elicited spoken narratives, although a much smaller sample, are basically in line with the more substantial written narrative skewings. Thus, out of a total 41 first-mention indefinites, we have the following breakdown (disregarding the subject:nonsubject dichotomy because of the numerically small scores):

Human	[+IS] = 91.7% (11/12)	Inanimate	[+IS] = 3.4% (1/29)
	[-IS] = 8.3% (1/12)		[-IS] = 96.6% (28/29)

[15]Referents which persist as active arguments until the end of their respective stories have been given a value of 30 mentions, this being roughly the average number callibrated for the first 20 type-tokens in the corpus. Corresponding tallies for the 41 first-mention indefinites in the oral narratives are:

Human	[+IS] = 14.0	Inanimate	[+IS] = 2.0
	[-IS] = 6.0		[-IS] = 1.3

[16]Cf. Chafe's (1980a:18) remark that "the backbone of a narrative consists of the introduction of people, their description, and especially their engagement in activities

which are worth telling about." Notice too that these claims concerning the linkage between humanness and IS-marking on first mention are at variance with Hopper and Thompson's (1983:11) contention that "the lexical semantic facts about nouns and verbs are secondary to their discourse roles (original emphasis)", although both approaches suffer from a certain amount of circularity.

[17]As far as I am aware, however, Hausa does not correspond to English in permitting the IS to particularize a nonidentifiable proper name as in 'A (certain) Mr. Smith came to see you yesterday.'

[18]For the purposes of this dissertation, I am including in the subject category arguments which are the grammatical subjects of object clauses, e.g. wani 'someone' in (21c).

[19]Since it would have been perfectly possible to insert the existential predicator akwai at the beginning of (22a), I am treating the presentative NP 'ya daya 'one daughter' as a subject argument.

[20]Notice, incidentally, that even though some of the various first-mention indefinites in (23) are not IS-marked, they are still all specified with some kind of modifier.

[21]A separate investigation of prepositional phrases is beyond the scope of the present study, although it might produce the kind of interesting distributional differences noted for German in Zubin (1979), a paper brought to my attention by Robert Kirsner (p.c.).

[22]This is not to say that that the IS and other modifiers are in complementary distribution -- they may co-occur, as example (40) below shows.

[23]Examples of IS-marked indefinite kin referents are available, however. In Imam's (1966) final volume, for instance, we encounter the following:

(36) a. An yi **wani** sarki a Garun Gabas,
IMPERS-PFV do IS emir-DO in Garun Gabas-ADV
b. wanda Ø Ø-ke da **wata** 'ya,
REL SUBJ-Ø CONC-Ø-IMPFV with IS daughter-ASSOC
c. a-na kira-n-ta Mama...
IMPERS-IMPFV call-of-her-VN Mama
(Imam 1966:34)

'There was a (certain) emir in Garun Gabas who had a

(certain) daughter by the name of Mama...'

where the associative-possessee indefinite kin-term wata 'ya 'a (certain) daughter' in (36b) is individuated via IS-marking.

[24]Russell Schuh (p.c.) has also reminded me that if the two NP's under consideration were assigned IS's, this would force the reading 'another wife of his', 'another daughter of his'. This additional disjunctive meaning of the IS is discussed in section 7 below.

[25]Proper name animals -- usually written with initial capitalized letters -- which regularly figure in animal-based folklore, e.g. Zaki '(the) Lion', Kura '(the) Hyena', Zomo '(the) Hare' etc., are never IS-marked on first mention, and so are excluded from the present analysis.

[26]The counts in Figures 2.1 and 2.2 do not include the 20 or so cases of the collocational wani abu 'something' -- mostly, but not exclusively, nonsubject forms -- encountered in the corpus.

[27]It is within the domain of nonsubject inanimate indefinites not coded by an IS on first mention, e.g. tsani (52c) and garma (53a), that one really encounters the interpretive problem noted earlier in 2.1, i.e. the potential ambiguity between a specific and nonspecific generic reading. The VP 'he was throwing a hoe' in (53a) would certainly allow a paraphrase like 'he was hoe-throwing' or some such conflation, but what would be a reasonably natural paraphrase of the VP 'he found a ladder' in (52c)? I think these kinds of problems support my earlier remarks about the difficulty of making a principled and workable distinction between the two interpretations.

[28]A typologically similar phenomenon is present in Sacapultec where, according to Du Bois (1981:279ff.), the same morpheme (xun) which introduces indefinites can also be used to code "topic switches". This topic switching function seems to differ from that of the Hausa Indefinite Specifier, however, in that a referent so marked signals only a definite interpretation -- it is glossed as 'the other' by Du Bois. In Hausa, this definite-identifiable disjunctive meaning is signaled by using a combination of the numeral 'one' and some form of the gender/number-sensitive definite determiner suffix. Fragments (55) and (56) illustrate this usage. In (55), Kalalatu has already eaten the wings of one of the two plump hens her husband has bought, and then decides to eat the wings of the other

remaining hen, saying:

(55) a. ...bari in cinye na
let-IMPERATIVE I-SUBJNCTV-SUBJ eat up of
ɗaya-r kuma
one-DET-DO too
b. don Ø su zo daidai...
so that SUBJ-Ø they-SUBJNCTV come equal

(Imam 1970:17)

'...let me eat up (the wings) of the other one (hen) too, so that they (the hens) are identical...'

In (55a) the NP ɗayar consists of the numeral ɗaya 'one' plus the feminine determiner suffix /-r/, and is used to mean 'the other one'. Kalalatu then devours one of the two hens completely and turns her attention to the remaining one:

(56) a. Sai Ø ta kama kaza-r guda
then SUBJ-Ø she-PFV seize hen-DET one-DO
b. Ø ta laɓume Ø.
SUBJ-Ø she-PFV devour DO-Ø

(Imam 1970:17)

'Then she seized the other hen and devoured (it).'

In this case, it is the direct object NP kazar 'the hen' which bears the same /-r/ suffix, followed by the quasi-numeral guda 'one'.

CHAPTER 3

CONDITIONS ON THE SELECTION OF ZERO ANAPHORA AND PRONOMINAL ANAPHORIC REFERENCE

1. Introduction

The main thrust of this chapter is an investigation of the two major categories of nonexplicit reference -- zero anaphora and pronoun anaphora -- and their exploitation in the coding of previously-mentioned third-person human, animal and inanimate referents. The chapter begins, in section 2, with some remarks on the corpus examined, and the methodological approach to the data. In section 3, I provide statistics and commentary on the distributional characteristics of all three major referential forms, i.e. nominal coreference (definite nouns, with or without accompanying modifiers, cf. 3.3), in addition to zero (3.1) and pronominal anaphora (3.2). This particular section thus functions as the quantitative and analytical starting point for the remainder of the dissertation.

Sections 4 and 5 address the key issues of the present chapter. In section 4, I attempt to account for the factors conditioning the selection of full pronouns (4.1) as opposed to zero anaphora (4.2) in the postverbal direct

object position, using the same salience:coding hypothesis developed in Chapter 2 to explain the distribution of IS-marking of indefinite nominals. And in section 5, we take a close look at the kinds of contexts which permit the omission of various types of item associated with the preverbal subject argument, including: zero anaphora of the subject itself (5.1); dropping of the subject concord-marker (5.2); and omission or replacement of the auxiliary element (5.3).

2. Data base and methodology

The written narrative data which form the baseline for the introductory overview given in section 3 are culled from the non-speech portions of the first five stories in Imam (1970:6-29) -- over 1,100 tokens of the three referential categories in question. And the spoken narrative material is from the four spontaneously-elicited texts at my disposal (cf. Chapter 1), containing almost 400 tokens in all.

In attempting to determine the ranges within which the various reference-types are typically employed, Chafe (1976) has suggested that the ongoing choices made by the encoder are related to his/her perception of the decoder's state of knowledge of a given referent at a given point in the communication (cf. too Clancy 1980). Chafe (1974, 1976, 1980a) has, moreover, made the additional important

point that "givenness" can gradually disappear from the decoder's consciousness with the simple passing of time, an observation which has received ample support from such studies as Kirsner (1979), Kirsner and van Heuven (1980), Bernardo (1980), Clancy (1980), and Givón (1983). And a number of psycholinguistic studies have independently demonstrated that the passage of time and/or the intervention of potentially ambiguous referents have an observable impact upon the decoder's ability to process incoming information (cf., inter alia, Norman 1969; Klatzky 1975). The general hypothesis emerging from these reports, therefore, states that when the cognitive factors of "time" and "interference", as defined below, combine to impair the decoder's ability to uniquely identify a given referent, then the encoder will compensate by using a more explicit referential form, e.g. a full nominal in preference to a "leaner" pronominal form.

Turning now to the methodological details, the impact of the "time" dimension on referential choice was assessed in terms of the linguistic unit of the "clause". This measure was taken by calculating the number of distinct clause boundaries separating two successive mentions of a referent.[1] In other words, if reference to a particular argument in one clause -- whatever the form used -- was followed by an additional mention of the same referent in the immediately following clause, then this was counted as

one intervening clause boundary for the purposes of the scores.[2] And the factor of "referent interference" was objectified by counting the number of references to maximally similar arguments intervening between two successive mentions of a given referent, i.e. "potentially" confusing third-person (singular/plural) arguments identical with regard to semantic class, e.g. human/nonhuman, and number-gender specification.

3. Distribution of the three referential classes: an overview

This section is designed to provide the reader with an introductory profile of the three categories now illustrated in sentences (1-3):

(1)	**Yara-n** suka zo. **boys-DET-SUBJ** they-PFV come '**The boys** came'		NP coreference
(2) a.	Yaran suka same **shi**. boys-DET-SUBJ they-PFV find **him-DO PRO** 'The boys found **him**'		Direct object pronoun anaphora (cf. 4.1)
b.	Yaran suka samu **ø**. boys-DET-SUBJ they-PFV find **DO-ø** 'The boys found **ø** (it).'		Direct object zero anaphora (cf. 4.2)
(3) a.	**ø** Suka zo. **SUBJ-ø** they-PFV come '**ø** They (the boys) came.'		Zero anaphora of the subject (cf. 5.1)

This chapter is largely concerned with the referential competition between the strategies exemplified in (2a-b), i.e. pronominal versus zero anaphoric reference in the postverbal direct object position, in addition to the conditions governing the acceptability of preverbal subject zero anaphora as illustrated in (3). Table 3.1 now summarizes data on the relative cross-text frequencies of the categories in question.

TABLE 3.1

Summaries of Relative Frequencies of Referential Forms in Written and Spoken Narratives

	Written Narratives		Spoken Narratives	
Referential Type	No.	%	No.	%
Zero anaphora	636	54.3	200	53.6
Pronominal anaphora	221	18.9	88	23.6
Nominal coreference	315	26.8	85	22.8
Totals	1172		373	

Inspection of the statistics in Table 3.1 reveals that the relative skewings computed for the three referential categories display a marked similarity in both the written and

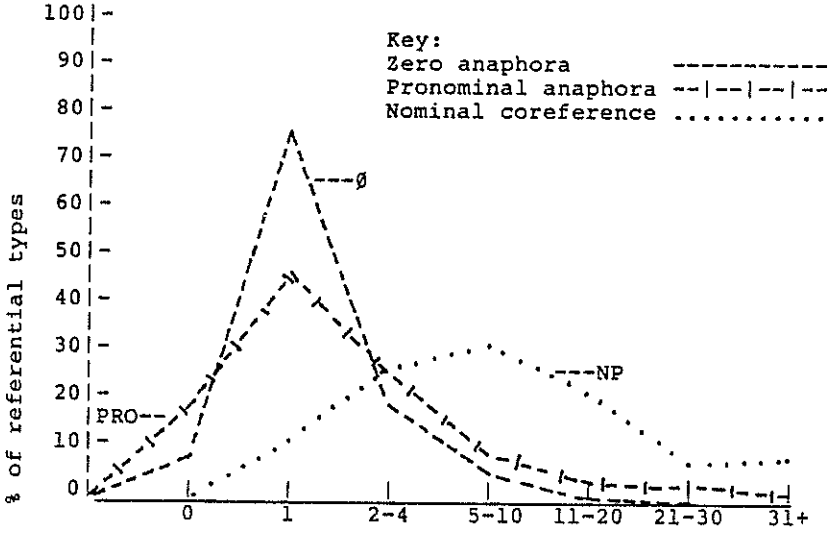
spoken texts. The data also demonstrate that zero anaphora constitutes the majority of the ongoing referential choices made -- 54.3% and 53.6% respectively -- a higher percentage, therefore, than the combined figures for the two other major categories -- 18.9% and 23.6% respectively for pronoun anaphora, and 26.8% and 22.8% for nominal coreference. Comparing these findings with those reported by Clancy (1980:132ff.) for the English and Japanese Pear Film accounts, we may note the following facts: Hausa (54.3%/53.6%) is somewhat intermediate between Japanese (73.2%) and English (20.5%) with regard to the relative frequency of ellipsis; English (63.3%) is characterized by a higher incidence of anaphoric pronominalization;[3] and with respect to the occurrence of nominal coreference, Hausa (26.8%/22.8%) approximates Japanese (26.8%), and has a slight edge over English (15.7%).

Having looked at a representative sampling of the relative frequencies of the referential options, we now proceed to the task of evaluating the choices made in terms of the cognitive constraints of time and interference. Figures 3.1-3.4 (cf. Appendices III-VI) now furnish relevant data on the interface between these two constraints and referential choice. In these figures, the percentage values on the vertical axis specify the percentage occurrences of a referential category in accordance with the numerical limits specified on the horizontal

axis.

Figure 3.1. **Distribution of the Three Referential Categories with respect to the Number of Intervening Clauses.**

Written Narratives



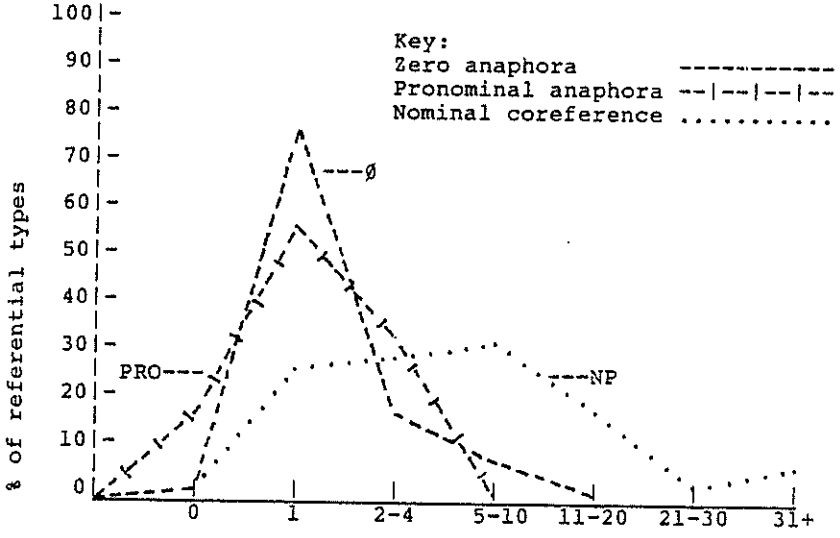
Number of distinct clause boundaries separating two successive mentions of the same referent

Average no. of intervening clauses:

Zero anaphora:	1.5
Pronominal anaphora:	2.4
Nominal coreference:	13.9

Figure 3.2. **Distribution of the Three Referential Categories with respect to the Number of Intervening Clauses**

Spoken Narratives

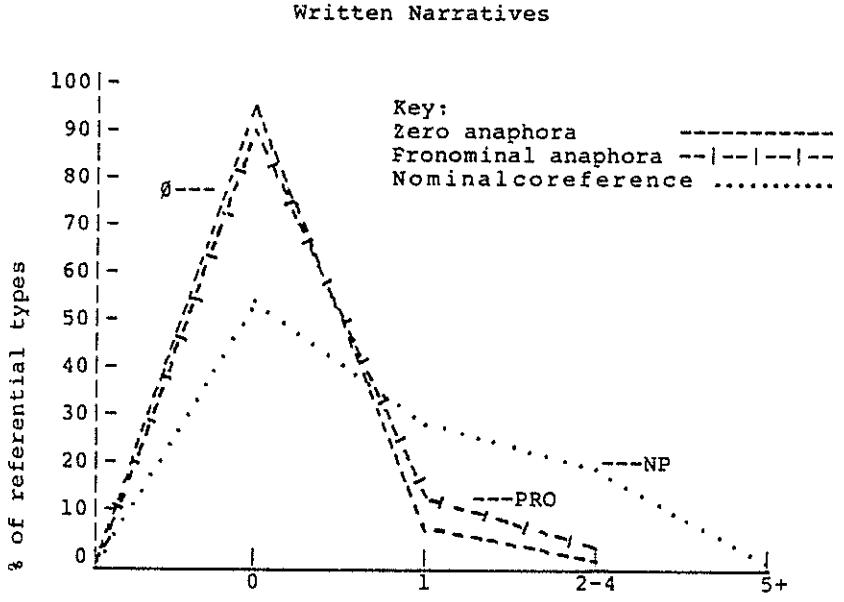


Number of distinct clause boundaries separating two successive mentions of the same referent

Average number of intervening clauses:

Zero anaphora:	1.8
Pronominal anaphora:	1.5
Nominal coreference:	7.7

Figure 3.3. Distribution of the Three Referential Categories with respect to the Number of Intervening Confusers

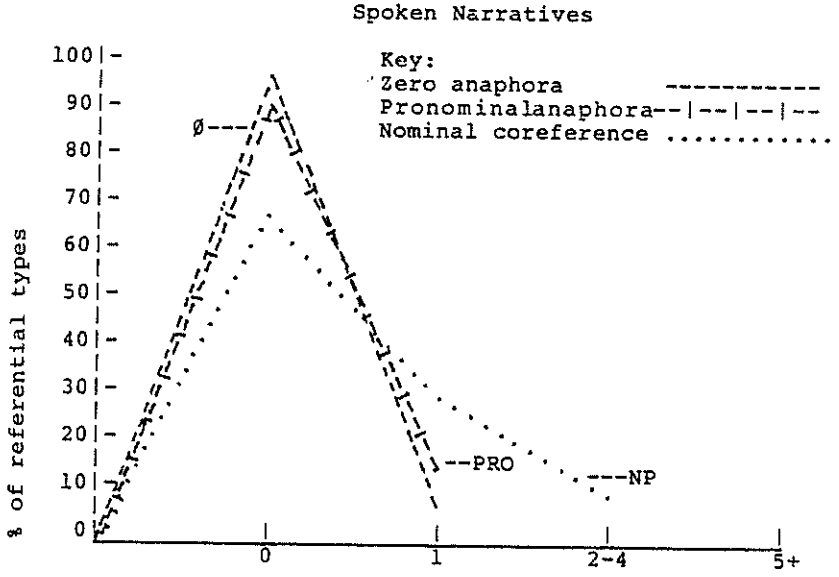


Number of maximally similar referents occurring between two successive mentions of the same referent

Average number of intervening referents:

Zero anaphora:	0.1
Pronominal anaphora:	0.1
Nominal coreference:	0.8

Figure 3.4. Distribution of the Three Referential Categories with respect to the Number of Intervening Confusers



Number of maximally similar referents occurring between two successive mentions of the same referent

Average number of intervening referents:

Zero anaphora:	0.1
Pronominal anaphora:	0.1
Nominal coreference:	0.5

The information summarized in Figures 3.1-3.4 point to several major patterns which are basically consistent with

the hypothesis concerning the interpredictability between discourse (dis)continuity -- as measured in terms of time and interference -- and exploitation of an attenuate or explicit referential mechanism. Taking each category in turn, the discernible trends are as follows:

3.1. Zero anaphoric reference

Figures 3.1 and 3.2 (cf. Appendices III-IV) reveal that the vast majority of zero anaphor tokens occurred in the tightly restricted context of the clause immediately following the last discourse control of the same referent -- a remarkably similar 81.0% and 76.5% for the written and spoken narratives respectively -- with 15.3%/16.0% following 2-4 intervening clauses. The average number of intervening clause boundaries is 1.5 and 1.8 respectively for the two narrative-forms. Turning to Figures 3.3 and 3.4 (cf. Appendices V-VI), we may note that almost all the recorded cases of zero anaphora occurred in environments where no intervening maximally similar referent was present -- 92.9%/94.5% -- with an identical median of 0.1 potentially ambiguous confusers computed for both genres. Sections 4.2 and 5.1 take a detailed look at direct object zero anaphora and subject zero anaphora respectively.

3.2. Pronominal anaphora

As with zero anaphora, most instances of pronoun anaphora were also encountered in discourse environments where only one clause boundary separated two successive mentions of a referent -- 51.1% and 53.4% respectively, though pronouns are less well represented in this group than zero anaphors. 27.6% and 30.7% of pronouns occurred after a gap of between 2-4 clauses, and Figures 3.1 and 3.2 also show that 9.9% and 14.8% were encountered within the same clause as the antecedent control. The 35 tokens in this group (cf. Appendices III-IV) were mainly suffixal possessive pronouns, occurring in such contexts as 'They$_j$ left their$_j$ camp', plus a handful of indirect object pronouns, e.g. 'His$_j$ friends laughed at him$_j$'. The average number of intervening clauses for the written and spoken texts is 2.4 and 1.5 respectively, i.e. not significantly higher then the 1.5/1.8 averages computed for the zero anaphora strategy. Figures 3.3 and 3.4 also demonstrate that the interference/confuser percentages are in fact the same as those for zero anaphora -- means of 0.1 for both the written and spoken texts, with 88.7% and 90.9% of all pronouns occurring without any potential confusers in the vicinity. Section 4.1 discusses direct object pronominal anaphora in depth.

3.3. Nominal coreference

The data in Figures 3.1 and 3.2 reveal that the largest percentages of coreferential nominal forms occurred when between 5-10 clauses had intervened -- 32.4% and 29.3% for written and spoken texts, followed by averages of 26.7% and 27.1% for the 2-4 clause grouping. In the written texts, only 8.8% of NP's fell into the 1-clause boundary category, and at the other extreme, far higher proportions of coreferential NP's were selected over discourse stretches involving 11-20, 21-30, and 30+ clauses. And the average number of intervening clauses for NP coreference was 13.9 and 7.7 respectively, i.e. significantly higher than the averages computed for the nonexplicit pronoun and zero categories. Figures 3.3 and 3.4 also reveal a discernible increase in the average number of potential confusers present between a nominal form and its immediately preceding antecedent -- in both narrative-types, 28.3% occurring with a single intervening competitor, 17.5% and 7.1% with between 2-4 potential confusers, and with overall averages of 0.8 and 0.5 respectively. Chapters 4 and 5 investigate the discourse behavior of the various combinations of nominal coreference, i.e. nouns with or without accessory modifiers of various types.

The foregoing material and observations were intended to provide an analytical point of departure for the rest

of the dissertation. The remainder of this chapter is now devoted to a detailed investigation of the two least explicit referential forms -- zero and pronominal anaphora.

4. Anaphoric reference to the direct object

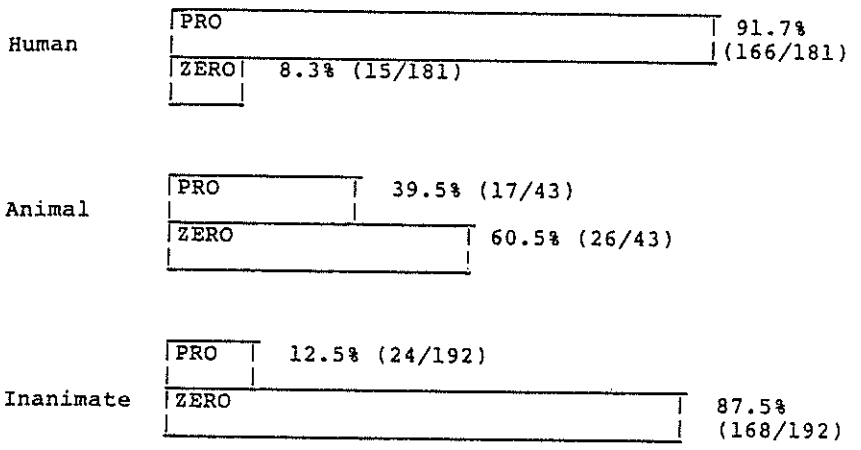
Postverbal direct object arguments allow two possible anaphoric options -- a full pronoun or zero anaphora.[4] Hausaists have had little to say on the competition between these two referential options. Parsons (1960:20), for example, writes that "the omission of this object word after the verb is optional only", and that "there is a good deal of stylistic latitude with regard to the repetition, or non-repetition, of an object with a series of transitive verbs." Cowan and Schuh (1976:135) are a little more explicit and state that "the direct object pronoun is optional with most verbs when you are referring to a previously mentioned specific, concrete object [original emphasis]", adding that "the pronoun is preferably not used when referring to an inanimate object. It is normally used when referring to a human." And finally, in Jaggar (1983:379), I claim that "subsequent to full first mention, a given NP referent may then admit of either zero or pronominal anaphoric recall in all clauses excepting the sequence-final one, where ellipsis is preferred; at

the same time, however, there is a marked predilection for picking up [+human] referents with a pronominal device, again in all but the final clause where a zero strategy is most commonly encountered." In this section, I will attempt to account for the formal distribution of full pronoun vs. zero anaphora of the direct object in terms of the salience:coding hypothesis.

Recall that, in Chapter 2, it was demonstrated that human indefinites were assigned a significantly higher proportion of Indefinite Specifiers than were nonhumans, and that this IS-attraction was a function of the greater discourse salience of human participants, defined as the degree to which a referent is deployed as an autonomous argument within the text, and measured in terms of the number of discourse mentions. Within the domain of inexplicit definite reference now under consideration, the same hypothesis will predict the following:

(5) Highly-salient, highly-deployable human referents should correlate with the overt marking option, i.e. DO pronoun anaphora; conversely, inanimates should favor nonovert DO zero anaphora, with animals perhaps representing an intermediate category.

The distributional statistics in Figures 3.5 and 3.6 below basically substantiate these predictions (cf. Appendices VII-VIII):[5]



Data source: Imam 1970:6-93.

Figure 3.5 **Distribution within Written Narratives of Direct Object Zero Anaphora and Direct Object Pronoun Anaphora with respect to Semantic Class.**

Figure 3.6 looks at the same data from a different perspective:

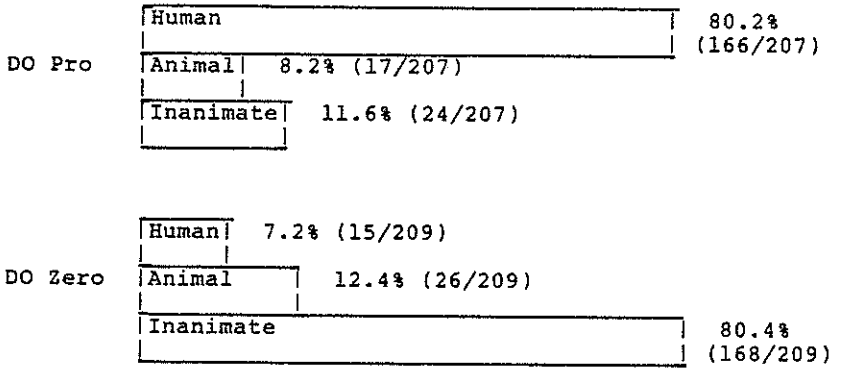


Figure 3.6 **Distribution of Direct Object Pronouns and Direct Object Zero Anaphors in Written Narratives: Numbers and Percentages of Totals according to Semantic Class.**

Even a cursory inspection of the cross-text data summarized in Figures 3.5 and 3.6 reveals some compelling distributional patterns which are consistent with the claims of the salience:coding hypothesis. Thus, just as highly-deployable human participants display a strong statistical preference for IS-marking on first mention, so too is there a marked tendency to anaphorically recall these same high-ranking human referents with the "heavier" of the two referential options now under consideration, i.e. a full direct object pronoun in preference to zero anaphora.[6] These findings are of additional significance,

moreover, since they provide independent support for the claim advanced in Chapter 2 that it is not simply the syntactic status of a given argument which influences morphological coding; rather, it is the intersection of syntactic function and semantic class membership which is the determining factor.[7]

Before proceeding with the analysis, I would like to call the reader's attention to a possible alternative explanation of the variant distributions noted in Figures 3.5-3.6. Taking, for the sake of argument, the two extreme categories of humans and inanimates, their highly contrastive distributions might be explicable in the following discourse terms: human referents are encoded with a much greater proportion of direct object pronouns (91.7%) than inanimates (12.5%) because they are found in contexts where the writer judges that the addressee would have difficulty assigning unique and proper identity to the referent if the less explicit strategy of zero anaphora were used, usually because too much "time" has elapsed -- measured, as ever, in terms of the number of distinct clauses boundaries intervening between two successive mentions of the same referent. The data now presented in Table 3.2 effectively rule out any such account, however.

TABLE 3.2

Distribution of Direct Object Pronouns and Direct Object Zero Anaphora with respect to Time and Interference.

Anaphoric Type	Written Narratives		Spoken Narratives	
	Tokens	Average No. of Clauses/ Confusers	Tokens	Average No. of Clauses/ Confusers
DO Pro	207	2.0/0.07	27	1.6/0.04
DO Zero	209	1.8/0.05	22	1.9/0.1

Table 3.2 shows that the average computed "look-back" to the immediately preceding discourse mention for direct object pronouns, the vast majority of which denote human referents (80.2%), is 2.0 clauses in the written stories, and that this value is only fractionally higher than the average 1.8 computed for zero anaphors in the same direct object position, the great majority of which (80.4%) have inanimate referents; and in the oral narratives, Table 3.2 shows that the score for the pronoun category is actually slightly less than the average calculated for zero anaphora -- 1.6 vs. 1.9 clauses. In other words, the competition between full pronoun and zero anaphora, unlike the choice of explicit nominal vs. nonexplicit pronoun or zero

for instance, cannot be attributed to cognitive constraints relating to limitations on short-term memory capacity; rather, in order to account for the numerical disproportions summarized in Figures 3.5-3.6, we need to consider, once again, the factor of referent salience -- humans, as the most prominent discourse participants, are generally assigned heavier morphology than their less salient nonhuman counterparts.

Let us now take a close look at the findings, beginning with the option of pronoun anaphora in the direct object position.

4.1. Direct object pronominal anaphora

The data in Figures 3.5 and 3.6 show that, as anticipated, highly salient human participants were coded with a substantially greater proportion of direct object pronouns than both the animal and inanimate categories. In most of the following citations, identification of the antecedent arguments is relatively transparent, and so I shall only comment on cases where such identification is not immediately apparent, or where unusual features are in evidence.

4.1.1. Human referents [+ DO PRO]

A striking 91.7% (166/181) of anaphoric reference to human arguments in the direct object position were full

pronouns. Excerpts (6-8) are illustrative:

(6) a. Sai Ø ya kama hannu-n-sa,
then SUBJ-Ø he-PFV seize hand-of-his-DO
b. Ø zai wuce da shi ciki-n
SUBJ-Ø he-FUT pass by with him-ASSOC inside-of
gida
house-LOC
c. Ø ya je
SUBJ-Ø he-SUBJCTV go
d. Ø ya tuhunce **shi**,
SUBJ-Ø he-SUBJCTV interrogate **him**-DO PRO
e. in Ø ya ki fadi-n gaskiya
if SUBJ-Ø he-PFV refuse state-of-VN truth
f. Ø ya yanka **shi**...
SUBJ-Ø he-SUBJCTV execute **him**-DO PRO

(Imam 1970:46)

'Then he (the emir) seized his (the stable boy's) hand, and was about to go by with him into the house to go and interrogate **him**, and if he refused to tell the truth he would execute **him**...'

(7) a. Wani dogari mai kili ya sake
IS bodyguard with horse-SUBJ he-PFV release
shi
it-DO PRO
b. Ø ya bi **shi**,
SUBJ-Ø he-PFV follow **him**-DO PRO
c. Ø ya sa doki
SUBJ-Ø he-PFV put horse-DO
d. Ø ya banke **shi**.
SUBJ-Ø he-PFV knock over **him**-DO PRO

(Imam 1970:41)

'A (certain) bodyguard with a horse gave it rein and followed **him** (the thief), and he set the horse and knocked **him** over.'

Fragment (8) is from one of the Pear Stories. The speaker has been talking about how fortunate it was for the 'mango man' that some boys were able to help him following his fall:

(8) a. Shikenan sai Ø ya yi sa'a,
OK then SUBJ-Ø he-PFV do luck
b. ga wasu yara sun zo daidai
PRESENT IS boys-SUBJ they-PFV come exactly
guri-n,
place-DET-LOC
c. Ø su-na wasa.
SUBJ-Ø they-IMPFV play
d. Sai yara-n suka zo
then boys-DET-SUBJ they-PFV come
e. Ø suka taimake **shi**...
SUBJ-Ø they-PFV help **him**-DO PRO

(Speaker 2)

'OK he (the mango man) was lucky, there were some boys, they had come right to the place and were playing around. Then the boys came and helped **him**...'

All the above cases of pronoun exploitation are typical in that prior reference is only a short distance to the left, and no maximally similar arguments are present to compete for reference.

4.1.2. Animal referents [+DO PRO]

Figures 3.5 and 3.6 reveal that this subcategory is in the minority, with a percentage of 39.5% (17/43) -- closer in fact to the 12.5% calculated for inanimates than it is to the 91.7% human value. Fragment (9) is illustrative:

(9) a. Da Ø suka jima da ita
when SUBJ-Ø they-PFV spend time with her-ASSOC
a hannu,
in hand-LOC
b. sarki ya ce
emir-SUBJ he-PFV say

c. ø su sake ta.
SUBJ-ø they-SUBJNCTV release her-DO PRO
d. ø Suka sake ta,
SUBJ-ø they-PFV release her-DO PRO
e. ø ta tashi.
SUBJ-ø she-PFV fly off

(Imam 1970:8)

'When they (the brothers) had spent some time holding her (the heron), the emir said they should release her. They released her and she flew off.'

4.1.3. Inanimate referents [+DO PRO]

In keeping with the predictions of the hypothesis, inanimate referents have by far the lowest percentage of full pronoun tokens -- 12.5% (24/192) of the total count. Fragment (10) exemplifies the subcategory:

(10) a. ...sai ø ya samo dogwaye-n sanduna
then SUBJ-ø he-PFV get long-of sticks
guda shida,
unit six-DO
b. ø ya auna tsawo-n-su,
SUBJ-ø he-PFV measure length-of-them-DO
c. kowa ø-na kallo,
everyone-SUBJ CONC-ø-IMPFV watch-VN
d. ø ya yanke su daidai da
SUBJ-ø he-PFV cut them-DO PRO exactly with
juna.
each other-ASSOC
e. ø Ya ɗauka ø
SUBJ-ø he-PFV take DO-ø
f. ø ya rarraba musu ø...
SUBJ-ø he-PFV distribute to them-IO PRO DO-ø

(Imam 1970:28)

'...then he (the judge) got six long sticks and measured their lengths, everyone was watching, and he cut them identically. He took ø (them) and distributed ø (them) to them...'

The inanimate NP dogwayen sanduna guda shida 'six long sticks' in (10a) is referred to with the direct object pronoun su in (10d); notice, however, that the writer soon resorts to the prevailing pattern for inanimates (cf. 4.2.3 below), using two zero anaphors in clauses (10e-f). Consider also excerpt (11):

(11) a. ...ø ya sa
SUBJ-ø he-PFV cause
b. dogarawa su tambaye shi
police-SUBJ they-SUBJNCTV question him-DO PRO
c. sai ø ya nuna inda
until SUBJ-ø he-PFV show where
ø ya ɓoye su.
SUBJ-ø he-PFV hide them-DO PRO

(Imam 1970:28)

'...he (the judge) had the police interrogate him (the thief) until he showed where he had hidden them (the money).'

The third person plural direct object pronoun su 'them' in (10c) is somewhat unusual in that it looks back more than 60 clauses to its preceding discourse control -- a (plural) NP kuɗi 'money' -- although the fact that 'the money' in question is an important motivating prop in the story may have something to do with the selection of a full pronoun. One maximally similar "potentially" ambiguous referent is also present between the two successive mentions -- sanduna 'sticks' -- though the prior context is clearly sufficient to prevent any confusion as to the referent of the pronoun.

4.2. Direct object zero anaphora

The term "direct object-ø" (DO-ø) is used to denote the structural gap which results when the postverbal direct object argument of a transitive verb is omitted. The information necessary for inferring the correct referent for the missing argument is presumed to be recoverable both from the prior discourse context, real-world knowledge, and/or the lexicosemantic subcategorization of the transitive verb itself. In some respects, direct object-ø represents the structural analog of zero anaphoric reference to the preverbal subject argument as discussed in section 5.1, the only difference being that with subject-ø, information regarding the identity of the controlling subject argument is, with the sole exception of cases of "concord-drop" (cf. 5.2), copied onto the auxiliary element. Figures 3.5 and 3.6 show, moreover, that the cross-text frequency of direct object-ø is almost as high as that of direct object pronominal anaphora.

4.2.1. Human referents [+DO ZERO]

This subcategory is in a distinct minority -- only 8.3% (15/181) of the entire anaphor count for humans, compared with 60.5% for animals and a substantial 87.5% for inanimate referents. In considering the subclass of human referents encoded with zero anaphors, it is important to note one fact of immediate relevance: in a close-knit, uninterrupted succession of coordinate transitive clauses, where the direct object arguments are identical, zero anaphora of all but the first reference in the chain is the overwhelming norm for all referent categories, i.e. regardless of semantic class membership. Just as significantly for the present claims, moreover, it turns out that all 15 recorded cases of zero-coded direct object human referents occurred in precisely such tightly-organized environments. Fragments (12-13) are illustrative:

(12) a. ...ø suka tarad da shi kwance
SUBJ-ø they-PFV find him-ASSOC lie-STAT
matacce.
dead
b. ø Suka kinkima ø
SUBJ-ø they-PFV carry DO-ø
c. ø suka kai ø gida
SUBJ-ø they-PFV take DO-ø home-LOC
d. ø suka binne ø.
SUBJ-ø they-PFV bury DO-ø

(Imam 1970:38)

'...they found him lying dead. They carried ø (him), took ø (him) home and buried ø (him).'

(13) a. Barayi suka yiwo waje,
Thieves-SUBJ they-PFV come outside-LOC
b. aka bi su kama kama,
IMPERS-PFV follow them-DO PRO at full pelt-ADV
c. aka bubbuge ø
IMPERS-PFV beat DO-ø
d. aka daure ø.
IMPERS-PFV tie up DO-ø

(Imam 1970:40)

'The thieves came outside, and the people (one) followed them at full pelt, beat ø (them) and tied

Ø (them) up.'

The DO-Ø forms in (12b-d) and (13c-d) typify the situation in chaining environments -- each zero device is separated from its immediately preceding discourse control -- whether pronoun or zero -- by only a single (transitive) clause boundary.

An additional feature of human referents should now be noted which is of direct relevance to the salience:coding hypothesis. The marked tendency to employ zero anaphora in tightly-organized transitive clause chains, as in (12-13), may in fact be overridden, but the only counter-examples evident in the corpus examined involved humans, e.g.,

(14) a. ...Ø ya ce
SUBJ-Ø he-PFV say
b. Ø ya-na da jika
SUBJ-Ø he-IMPFV with grandson-ASSOC
c. wanda ba abi-n da za a yi
REL NEG EXIST thing-DET REL FUT IMPERS do
masa
to him-IO PRO
d. Ø ya kadu.
SUBJ-Ø he-SUBJCTV be afraid.
e. In sarki ya-na so
if emir-SUBJ he-IMPFV want-VN
f. Ø ya gani,
SUBJ-Ø he-SUBJCTV see,
g. Ø ya kirawo **shi**
SUBJ-Ø he-SUBJCTV call **him-DO PRO**
h. Ø ya gwada **shi.**
SUBJ-Ø he-SUBJCTV test **him-DO PRO**

(Imam 1970:21)

'...he (the grandfather) said he had a grandson who could not be frightened. If the emir wanted to see,

he should summon **him** and test **him.**'

It is also worth pointing out that use of a full (direct object) pronoun, particularly in such contexts as (14h), represents a violation of Givón's (1983:18) claim that "the more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it" (original emphasis). The "topic" of (14h) is 'the grandson', last mentioned with a full pronoun in the preceding clause (14g) -- it could hardly be less "disruptive, surprising, discontinuous or hard to process etc.", yet it is coded with a full pronoun, not zero anaphora, in (14h).

4.2.2. Animal referents [+DO ZERO]

Almost twice as many animal referents were encoded by zero anaphors -- 60.5% (26/43) -- as were picked up by full pronouns -- 39.5% (17/43) -- in the direct object position. Extracts (15-17) exemplify this subcategory:

(15) a. ...Ø ya ce wa Kalalatu
SUBJ-Ø he-PFV say to Kalalatu-IO
b. Ø ta soye kaji-n nan
SUBJ-Ø she-SUBJCTV fry hens-DET DEM-DO
gaba daya.
all at once-ADV
c. Ø Ta tashi,
SUBJ-Ø she-PFV get up
d. baya-n Ø ya fige Ø
after-of SUBJ-Ø he-PFV pluck DO-Ø
e. Ø ya gyara mata Ø...
SUBJ-Ø he-PFV clean for her-IO PRO DO-Ø

(Imam 1970:16)

'...he told Kalalatu to fry those hens all at once. She (Kalalatu) got up, and after he had plucked ∅ (them) and cleaned ∅ (them) for her...'

(16) a. Ko da kyanwa ta ga ɓera
as soon as cat-SUBJ she-PFV see mouse-DO
b. sai ∅ ta yi tsalle
then SUBJ-∅ she-PFV do jumping
c. ∅ ta kama shi
SUBJ-∅ she-PFV seize him-DO PRO
d. ∅ ta kashe ∅
SUBJ-∅ she-PFV kill DO-∅
e. ∅ ta cinye ∅.
SUBJ-∅ she-PFV eat up DO-∅

(Imam 1970:51)

'As soon as the cat saw the mouse she jumped up, seized him, killed ∅ (him) and ate ∅ (him) up.'

Finally, excerpt (17) illustrates a case of DO-∅ which is noteworthy in that it looks back over an unusually long discourse stretch to its antecedent. The background is that a young man -- 'Halilu-the-Crafty' -- is attempting to indentify a particular heron, but is unable to do so:

(17) a. ∅ Ya nema ∅,
SUBJ-∅ he-PFV look for DO-∅
b. ∅ bai gane ta ba...
SUBJ-∅ NEG-he-PFV recognize it-DO PRO NEG

(Imam 1970:9)

'He looked for ∅ (the heron), (but) didn't recognize it...'

The preceding control for the direct object-∅ device in (17a) -- a direct object pronoun referring to the same

'heron' -- is a full 21 clauses to the left, though there are no maximally similar referents around which might compete for reference.[8]

4.2.3. Inanimate referents [+DO ZERO]

Figure 3.5 shows that, as per the hypothesis, zero anaphora was the overwhelming choice for inanimates occupying the direct object position -- a substantial 87.5% (168/192) of all inanimates being so coded. Fragments (18-22) are illustrative:

(18) a. ...∅ ya yi ta nema-n ƙaguwa
SUBJ-∅ he-PFV continue seeking-of-VN crabs
b. ∅ ya-na kama-wa ∅
SUBJ-∅ he-IMPFV catch-VN DO-∅
c. ∅ ya-na zuba-wa ∅ ciki-n
SUBJ-∅ he-IMPFV pour-VN DO-∅ inside-of
sanho.
grass bag-LOC
d. Da ∅ ya ' cika shi fal da
when SUBJ-∅ he-PFV fill it-DO PRO full with
ƙaguwa
crabs-ASSOC
e. sai ∅ ya ɗauko ∅
then SUBJ-∅ he-PFV take DO-∅
f. ∅ ya kawo ∅ gida
SUBJ-∅ he-PFV bring DO-∅ home-LOC
g. ∅ ya ɓoye ∅.
SUBJ-∅ he-PFV hide DO-∅

(Imam 1970:84)

'...he continued to look for crabs, catching ∅ (them) and pouring ∅ (them) into a grass bag. When he had filled it full with crabs, he took ∅ (it), brought ∅ (it) home, and hid ∅ (it).'

Extract (18) in fact exemplifies direct object-∅ of both

animals (18b-c) and inanimates (18e-g).

(19) a. ...ø ya yi ta miƙo musu
SUBJ-ø he-PFV continue hand to them-IO PRO
kaya,
goods-DO
b. ø su-na karɓa ø,
SUBJ-ø they-IMPFV receive-VN DO-ø,
c. ø su-na ajiye-wa ø.
SUBJ-ø they-IMPFV put down-VN DO-ø.
d. ø Ya zare zobe-n
SUBJ-ø he-PFV take off ring-DET-DO
e. ø ya ɓoye ø a kubaka-r
SUBJ-ø he-PFV hide DO-ø in hem-of
wando-n-sa...
trousers-of-his-LOC

(Imam 1970:90)

'...he continued passing out the goods to them, and they were receiving ø (them) and putting ø (them) down. He took off the ring and hid ø (it) in the hem of his trousers...'

(20) a. Da ø ya ga
when SUBJ-ø he-PFV see
b. ø sun yi yawa haka
SUBJ-ø they-PFV do plenty so
c. sai ø ya ji tsoro-n
then SUBJ-ø he-PFV feel fear-of
d. kada ø ya bar su
lest SUBJ-ø he-SUBJNCTV leave them-DO PRO
a bukka-r-sa,
in grass hut-of-his-LOC
e. wani ya zo
IS-SUBJ he-SUBJCTV come
f. ø ya sace ø.
SUBJ-ø he-SUBJCTV steal DO-ø

(Imam 1970:53)

'When he saw that they (the money) had become so plentiful, he was afraid lest he leave them in his grass hut and someone would come and steal ø (it).'

(21) a. Sai ø ta tsuguna
then SUBJ-ø she-PFV squat down

b. ø ta cinye ta sarai,
SUBJ-ø she-PFV eat up it-DO PRO completely-ADV
c. ø ta kwashi ƙasussuwa
SUBJ-ø she-PFV collect bones-DO
d. ø ta kai ø masai
SUBJ-ø she-PFV take DO-ø cesspit-LOC
e. ø ta zuba ø.
SUBJ-ø she-PFV pour DO-ø

(Imam 1970:17)

'Then she (the wife) squatted down, ate it (the chicken) up, collected the bones, took ø (them) to the cesspit and disposed of ø (them).'

The final example is from one of the Pear Film accounts:

(22) a. Da ø ya tsinko waɗannan 'ya-n
when SUBJ-ø he-PFV pick DEM DIMIN-of
itatuwa,
trees-DO
b. sai ø ya sanya ø a ciki-n
then SUBJ-ø he-PFV put DO-ø at inside-of
kwando.
basket-LOC
c. To sai can, sai da an jera kwanduna
OK then later after IMPERS-PFV arrange baskets
guda uku,
unit three-DO
d. sai ø ya-na nan,
then SUBJ-ø he-IMPFV there-LOC
e. ø ya-na ta tsinka ø
SUBJ-ø he-IMPFV continue pluck-VN DO-ø
f. ø ya-na ta zuba-wa ø
SUBJ-ø he-IMPFV continue pour-VN DO-ø
a ciki.
at inside-LOC

(Speaker 2-1)

'When he (the 'mango man') had picked those fruits he put ø (them) in a basket. OK later, after three baskets had been arranged in a row, there he was, continuing to pluck ø (them) and pouring ø (them) inside (the basket).'

5. Nonoccurrence of material associated with clausal subjects

In certain specifiable contexts, Hausa permits omission of the following items: the subject argument itself (subject-Ø, cf. section 5.1); the subject-concord morpheme (concord-drop, cf. 5.2); the preverbal auxiliary element (auxiliary-drop, cf. 5.3). Although these phenomena all involve some permutation (omission and/or replacement) of subject-related material, they nonetheless have different statuses -- only subject-Ø entails deletion of a full argument as such -- and so I have decided to refer to them jointly with the generalized description used in the section heading. Table 3.3 summarizes some comparative distributional data on these three devices with respect to the dimensions of time and interference.

TABLE 3.3

Distribution of Various Categories of Subject-Related Omission with respect to Time and Interference.

Item	Written Narratives		Spoken Narratives	
	Tokens	Average No. of Clauses/ Confusers	Tokens	Average No. of Clauses/ Confusers
Subject-Ø (5.1)	411	1.5/0.1	177	1.7/0.0
Concord-drop (5.2)	19	0.1	1	0.0
Auxiliary-drop (5.3)	6	0.3		

The statistics in Table 3.3 demonstrate that omission of these various subject-linked items is permissible only in highly continuous discourse contexts as measured in terms of the number of clauses intervening between prior discourse control and the device in question.[9] Let us now consider the categories in more detail.

5.1. Zero anaphora of the subject argument

Recall that a typical constituent configuration in Hausa -- with the subject argument overtly expressed -- would be as follows: subject NP + auxiliary + verb (+ ob-

ject NP), with the tense-aspect or mood of the verb encoded on the auxiliary element, which also copies the person, number, and gender of the preceding subject (cf. examples 1, 2a-b). And we have already seen numerous examples of this type of configuration in the foregoing text. Table 3.3 reveals that zero anaphoric reference to the subject (**subject-Ø**) is an extremely common phenomenon -- 411 and 177 tokens respectively for the written and spoken forms. Since, moreover, subject-Ø entails deletion of a full argument, it may be considered the structural analog of the direct object-Ø option (section 4.2), though the two phenomena do differ in that because the auxiliary element is usually inflected with a "pronoun-like" copy of the deleted subject, subject-Ø, unlike direct object-Ø, does incorporate information about the identity of the missing argument. The major distinguishing feature relates to the fact that whereas a zero:pronoun opposition is available for direct objects, there is no comparable opposition for subjects, i.e. subject pronouns do not co-occur with subject-concord. Comparison of the data in Table 3.3 with the figures given in Table 3.2 shows that the average intervening clause scores for subject-Ø are in fact remarkably close to the values calculated for both direct object-Ø and direct object pronoun anaphora -- averages of 1.5 and 1.7 intervening clauses for subject-Ø in the written and spoken texts, compared with means of

1.8 and 1.9 clauses respectively for direct object-Ø in the written and spoken texts, and 2.0 and 1.6 for full direct object pronoun anaphora. Fragments (23-27) provide further exemplification of the subject-Ø strategy:

(23) a. Isa Lamiri ya ɗebar masa
Isa Lamiri-SUBJ he-PFV collect for him-IO PRO
'ya-n sule goma,
DIMIN-of shilling ten-DO
b. Ø ya sallame shi.
SUBJ-Ø he-PFV dismiss him-DO PRO
c. Ø Ya yi ta kwasa-r saura-n
SUBJ-Ø he-PFV keep on take out-of remainder-DET-DO
d. Ø ya-na kai wa gida.
SUBJ-Ø he-IMPFV take to house-IO
e. Da Ø ya ƙulle su wuri guda
when SUBJ-Ø he-PFV hide them-DO PRO place one-LOC
f. sai Ø ya zauna,
then SUBJ-Ø he-PFV sit down
g. Ø ya yi godiya ga Allah bisa ga
SUBJ-Ø he-PFV do thanks to Allah-IO on
wannan baiwa da Ø ya yi masa.
DEM gift REL SUBJ-Ø he-PFV do to him-IO PRO

(Imam 1970:12)

'Isa Lamiri collected ten shillings for him and Ø dismissed him. Ø (He) continued to take out the rest (of the money) and Ø take it home. When Ø (he) had locked it up in one place (he) Ø sat down and Ø thanked Allah for this gift that Ø (he) had given him.'

Fragment (23) begins with mention of the full NP subject -- the proper name Isa Lamiri in clause (a). In the remaining clauses (b-g), however, this same subject argument is fully ellipted, though the information that the same referent persists as clause subject is preserved on the third-person singular masculine preverbal auxiliaries ya (perfective) and yana (imperfective). Notice too the

tight sequencing which is a common characteristic of subject-ø -- each occurrence of the device looks back across one clause boundary only to the previous control. This type of "stage-holding" is thus consonant with Givón's (1980:306) claim that "the maintenance of the same NP onward in the subject role...is taken to be the most predictable, most expected, least-surprising and least disruptive strategy in topic-NP identification" (cf. too Nichols 1981). Excerpt (24) illustrates further the close sequencing and lack of intervening confusers which typify this function -- in this case with a third-person plural referent. Three young men arrive at an emir's palace and are questioned by one of the emir's officials:

(24) a. ø Suka fadi sunaye-n-su ɗai ɗai,
SUBJ-ø they-PFV tell names-of-them-DO one one-ADV
b. ø suka gaya masa kuma
SUBJ-ø they-PFV tell to him-IO PRO and
c. ø sun fito daga Kona ne,
SUBJ-ø they-PFV come from Kona COP
d. ø za-su yawo-n duniya,
SUBJ-ø FUT-they wander-of world
e. don ko Allah ya sa
so that Allah-SUBJ he-SUBJCTV cause
f. ø su sami inda
SUBJ-ø they-SUBJCTV find where
g. ø za-su raba gardama-r da[10]
SUBJ-ø FUT-they resolve quarrel-DET REL
ø-ke tsakani-n-su.
CONC-ø-IMPFV between-of-them

(Imam 1970:7)

'ø (They) stated their names one by one, and ø told him that ø (they) had come from Kona, and ø were going to wander around the world, so that Allah could cause ø (them) to find a means by which ø (they) could resolve the differences between them.'

And at the beginning of one of the Pear Film accounts we encounter a similar tight chaining of subject-ø devices following upon the introduction of the 'mango man' into the story:

(25) a. Da farko dai wani mutum ne,
at first-ADV well IS man COP
b. ø ya je...
SUBJ-ø he-PFV go
c. ø ya-na tsinka-r mangwaro-n-shi.
SUBJ-ø he-IMPFV pluck-of-VN mangoes-of-his-DO
d. ø Ya-na da kwanduna guda uku.
SUBJ-ø he-IMPFV with baskets unit three-ASSOC
e. Shikenan ø ya cika kwando guda biyu,
That was that SUBJ-ø he-PFV fill basket unit two-DO
f. ø ya hau ka-n mangwaro-n,
SUBJ-ø he-PFV climb top-of mango tree-DET-LOC
g. ø zai je
SUBJ-ø he-FUT go
h. don ø ya tsinko mangwaro-n
in order to SUBJ-ø he-SUBJCTV pluck mangoes-DET
da ø zai cika kwando-n-shi na ƙarshe...
REL SUBJ-ø he-FUT fill basket-of-his of end-DO

(Speaker 1)

'Well, first of all it's a man, ø (he) has gone... ø (he) is picking his mangoes. ø (He) has three baskets. OK, ø (he) has filled two baskets, ø (he) has climbed the mango tree, and ø (he) is about to go ø to pick the mangoes (with) which ø (he) will fill his third basket.'

Discounting environments in which a block of direct speech quotations intervene between two mentions -- the impact of this variable is discussed in Chapter 5 -- the maximum gap I encountered for subject zero anaphora was in the region of 3-4 clause boundaries. Extract (26) provides exemplification:

(26) a. Baya-n kama-r kwana-n wata
After-of about-of day-of month
mata-r Isa ta yi ciki.
wife-of Isa-SUBJ she-PFV do stomach
b. Wata-n nan bai dawo ba
month-DET DEM-SUBJ NEG-it-PFV return NEG
c. sai da ø ta haifi ɗa namiji.
except that SUBJ-ø she-PFV bear son male-DO
d. ø Bai manta da abi-n da
SUBJ-ø NEG-he-PFV forget with thing-DET REL
aka gaya masa ba.
IMPERS-PFV tell to him-IO PRO NEG

(Imam 1970:14)

'After about a month Isa's wife became pregnant.
Within a year ø (she) had given birth to a son.
ø (He/Isa) didn't forget what he had been told...'

The antecedent for the subject-ø form in (26d) is Isa, designated by the proper name in (26a), i.e. 3 clauses to the left. Notice too that if we are to interpret the notion of "maximally similar potential confuser" quite literally, as I have in fact done for purposes of the counts, then the intervening NP ɗa namiji 'son' in (26c) should, strictly speaking, qualify as such. However, real-world knowledge -- a newly-born child could hardly control the cognitive verb 'forget' in (26d) -- plus the prior discourse context -- we, the readers, know that Isa has been given certain instructions in an earlier dream, render such an interpretation clearly implausible. Consider too example (27), again somewhat atypical for subject-ø with respect to the number of intervening clauses and potential confusers. A husband Kalala is under the impression that a visitor Baƙo has stolen one of his chick-

ens, and so runs after him, knife in hand, shouting that he wants him, i.e. Baƙo, to leave him one of the chickens at least. Baƙo thinks that he means to cut off one of his ears, however:

(27) a. Baƙo tsammani ya-ke
Baƙo-SUBJ thinking he-IMPFV
b. Kalala ø-na nufi-n
Kalala-SUBJ CONC-ø-IMPFV mean-of
c. kunne-n-sa ɗaya kaɗai ø ya-ke so
ear-of-his one only SUBJ-ø he-IMPFV want-VN
d. ø ya yanka.
SUBJ-ø he-SUBJCTV cut off
e. ø Ya waiwaya
SUBJ-ø he-PFV turn round
f. sai ø ya ga
then SUBJ-ø he-PFV see
g. ø ya taso masa da wuƙa...
SUBJ-ø he-PFV come to him-IO PRO with knife-ASSOC

(Imam 1970:19)

'Baƙo thought Kalala meant that ø (he/Kalala) wanted ø to just cut off one of his ears. ø (He/Baƙo) turned round and ø saw that ø (he/Kalala) had started for him with a knife...'

Our immediate concern is with how the referent Baƙo is traced, following full lexical mention as the subject of (27a). Baƙo re-emerges as the subject once again in (27e), following a switch of subjects to Kalala in (27b-d) -- notice that Baƙo is in fact picked up in (27c) with the possessive pronoun suffix -sa, which means that the (27e) reference has a look-back of 2 clauses. Technically speaking, we have a case of "potential" ambiguity again, with both Baƙo and Kalala competing for the subject zero anaphoric reference in (27e) onwards -- they are both

third-person masculine singular human arguments. However, the preceding discourse context provides a natural basis for inferring that it is Baƙo who is being chased by an irate, knife-wielding Kalala, and this guarantees that only the Baƙo referent could be interpreted as controlling the actions depicted by the verbs in (27e-f).

5.2. Subject concord-drop

The auxiliary element in Hausa breaks down into two morphological types: in some tense-aspects, e.g. the perfective, the subject agreement-morpheme and tense-aspect marker are fused, i.e. the auxiliary is synthetic; in others, e.g. the imperfective, the auxiliary is analyzable as agglutinative. When the auxiliary is of this latter polysynthetic type, i.e. where the concord and tense-aspect functions are encoded by discrete morphemes -- omission of the concord-marker copying the subject features is permissible in certain well-defined contexts, with the result that only the tense-aspect morpheme remains overtly expressed. This phenomenon -- referred to as **"concord-drop"** -- is characteristic of two (affirmative) tense-aspects: the imperfective (neutral or relative) and, marginally, the iterative-habitual.[11] Selection of the concord-drop option is subject to the following strict conditions: in the vast majority of cases -- example (30b)

below is a conspicuous exception -- the overt equi-subject argument is located immediately to the left of the auxiliary (cf. intervening clause counts of 0.1/0.0 in Table 3.3), either in the form of a noun, full independent pronoun expressing focus, e.g. ita Ø-ke ce-wa haka (she-IND PRO-SUBJ CONC-Ø-IMPFV say-VN this) 'She says/claims this', or antecedent to a relative clause. Consider example (29). The wife Kalalatu has been frying some chickens, and the aroma is beginning to affect her:

(29) a. Ø Ta ci gaba da suya,
SUBJ-Ø she-PFV continue with frying-VN
b. ƙanshi Ø-na jifa-r-ta...[12]
aroma-SUBJ CONC-Ø-IMPFV throw-of-her-VN

(Imam 1970:18)

'She carried on frying, the aroma Ø overpowering her...'

In (29b) we have a typical example of imperfective concord-drop where the antecedent subject argument is located immediately to the left, in this case the first mention nominal ƙanshi 'aroma'. Notice too that first mention of a given referent in no way precludes the use of the concord-drop strategy.

Fragment (30) is of some interest since it contains the only case I recorded of a pair of concord-drop strategies occurring in successive clauses:

(30) a. Kalala Ø-na can
Kalala-SUBJ CONC-Ø-IMPFV there-LOC

b. ∅-na fama da washi-n wuƙa...
CONC-∅-IMPFV struggle with sharpening-of-VN knife

(Imam 1970:18)

'Kalala ∅ was there, ∅ struggling to sharpen the knife...'

The first instance of imperfective concord-drop occurs in (30a), in the same clause as its antecedent subject control -- the proper name Kalala. In (30b) we encounter use of the same mechanism, referring to the identical subject argument, but this time with a look-back of one clause. Without this particular example, the average look-back for concord-drop would have been zero instead of the 0.1 clause average listed in Table 3.3.

A single example of imperfective concord-drop was found in one of the oral accounts, in which the story-teller is recalling the death and burial of Alu, a former northern Nigerian emir:

(31) a. ...har Alu ya sami ciwo,
until Alu-SUBJ he-PFV get illness-DO
b. ∅ ya kwanta,
SUBJ-∅ he-PFV lie down
c. ∅ ya rasu.
SUBJ-∅ he-PFV die
d. Kabari-n-sa ma ∅-na nan...
grave-of-his-SUBJ and CONC-∅-IMPFV there-LOC

(Speaker 3)

'...until Alu became ill, laid down, and died. And his grave ∅ is still there...'

Fragments (32-33) exemplify the two cases I encountered of concord-drop in the iterative-habitual:

(32) a. ...sai ∅ ya ɗauki saura-n 'ya-n
then SUBJ-∅ he-PFV take rest-of DIMIN-of
kuɗi-n-sa
money-of-his
b. ∅ ya tafi kanti
SUBJ-∅ he-PFV go store-LOC
c. ∅ ya sawo tufafi iri-n waɗanda
SUBJ-∅ he-PFV buy clothes-DO kind-of REL
d. Sarki ∅-kan shiga da dare...
emir-SUBJ CONC-∅-HABIT enter at night-ADV

(Imam 1970:44)

'...the he (the stable-boy) took the rest of the money, went to the store, and bought the kinds of clothes that the emir ∅ put on at night...'

(33) a. Sai baya-n kwana bakwai Sarki
then after-of day seven-ADV emir-SUBJ
∅-kan zo
CONC-∅-HABIT come
b. da shi da Waziri su
and he-IND PRO and vizier-SUBJ they-SUBJNCTV
tambaye shi...
ask him-DO PRO

(Imam 1970:169)

'Then after seven days the emir ∅ would come, and he and the vizier would question him (the prisoner)...'

Again, the acceptability-conditions on occurrence of habitual concord-drop are just as restrictive as those constraining the more frequently-encountered imperfective concord-drop strategy.

5.3. Auxiliary-drop

In this final section we take a brief look at several

operations, all of which involve some permutation of the preverbal auxiliary element. Since, moreover, the phenomena in question entail the dropping of the auxiliary, together with substitution of alternative material in some cases, I have decided, in the interests of convenience, to use the cover-term **"auxiliary-drop"** to describe them all. The operations considered are, in order of presentation: dropping of the entire auxiliary in stative constructions (5.3.1); and replacement of the auxiliary with a nonfinite verbonominal form in purpose and participial clauses (5.3.2). The intervening clause figures in Table 3.3 reveal that the various items subsumed under the term "auxiliary-drop" display a close coincidence with concord-drop in that they may only be used in highly restricted contexts.

5.3.1. Stative auxiliary-drop

In one context in the language -- with stative constructions where an imperfective auxiliary is used before a stative verbal form -- it is possible to omit the entire imperfective auxiliary without requiring changes elsewhere in the VP.[13] The cross-text frequency of **stative auxiliary-drop** is extremely low, with example (35) one of only two tokens recorded:

(35) a. Ø Ta ruga waje-n ƙofa-r
SUBJ-Ø she-PFV rush place-of doorway-of
zaure
entrance hut-LOC
b. Ø ta leƙa
SUBJ-Ø she-PFV peep
c. ko Ø ta hango miji-n
if SUBJ-Ø she-SUBJCTV spot husband-DET-SUBJ
Ø tafe da baƙi-n...
STAT-Ø come-STAT with strangers-DET-ASSOC

(Imam 1970:17)

'She (Kalalatu) rushed to the entrance hut doorway and peeped out (to see) if she might spot the husband Ø coming with the strangers...'

Clause (35c) exemplifies the strategy in question -- the dropping of the entire imperfective auxiliary in position before the stative verbal form tafe 'coming'. The more explicit expression corresponding to this nonovert device would be as follows: full auxiliary, i.e. ya-na tafe (he-IMPFV come-STAT), or auxiliary stripped of the concord-marker (cf. 5.2), leaving Ø-na tafe (CONC-Ø-IMPFV come-STAT). What distinguishes the dropping, on the one hand, of either the imperfective concord-marker or the entire imperfective stative auxiliary from, on the other hand, use of the full auxiliary is the fact that the first two devices can only be exploited in the tightest possible environments -- basically when the antecedent subject argument, e.g. mijin 'the husband' in (35c), is to be found nestling in position immediately to the left of the attenuated device.

5.3.2. Purpose and participial clauses

Both these clause-types involve a switch to a nonfinite verbonominal form, uninflected for tense-aspect and subject-agreement, as an alternative to the use of an auxiliary + finite verb. Whether one considers these options to be instances of a sentence-bound rule of "Equi-NP deletion" (Russell Schuh [p.c.]), or simply an alternative syntactic device, it is again the case that they may only be selected if the antecedent subject controls reference of some kind in the close vicinity. Fragment (36) contains an example of a verbonominal used to express a purposive intent on the part of the preceding subject argument.

(36) a. ...ko kuwa Ø ta ce
or else SUBJ-Ø she-SUBJCTV say
b. Ø ta sauke Ø
SUBJ-Ø she-PFV put down DO-Ø
c. Ø ta shiga ɗaki **ɗauko**
SUBJ-Ø she-PFV enter hut-LOC **fetch-VN-PURPOSE**
kwano...
bowl-DO

(Imam 1970:16)

'...or she might say that she had put Ø (the chicken) down and gone into the hut to **fetch** a bowl...'

Clause (36c) exemplifies the purpose verbonominal in question -- ɗauko 'to fetch', itself paraphraseable with a subjunctive auxiliary + finite verb ta ɗauko 'so that she might fetch'.

Extract (37) illustrates the participial use of a nonfinite verbonominal form. Isa, the hero of the piece, is seated in an entrance porch when 'a Tuareg' comes in -- just as predicted in his dream of the previous day -- and promptly attacks him:

(37) a. ...Ø bai ce masa kanzil ba
SUBJ-Ø NEG-he-PFV say to him-IO PRO word NEG
b. sai Ø ya tsaya bisa ka-n Isa.
then SUBJ-Ø he-PFV stand over head-of Isa
c. **Da gani-n-sa**
with seeing-of-him-VN-PARTICIPLE
d. sai Isa ya tuna da mafarki-n
thenIsa-SUBJ he-PFV remember with dream-of
jiya...
yesterday-ASSOC

(Imam 1970:12)

'...he (the Tuareg) said nothing to him, then stood over Isa. **On seeing him,** Isa remembered yesterday's dream...'

It is worth mentioning that because the identity of the subject (and object) of the verbonominal participial phrase **Da ganinsa** 'On seeing him' in (37c) is not immediately obvious -- both masculine singular human referents Isa and 'the Tuareg' are in direct competition -- the writer feels it necessary to provide clarification by using a full NP subject in the following clause (37d), i.e. the proper name Isa. Finally, it is of interest to note too that if the writer had elected to use a full auxiliary + finite verb construction in this particular context, i.e. Da Ø ya gan shi sai Isa... (when SUBJ-Ø he-

PFV see him-DO PRO then Isa-SUBJ) 'When he saw him, Isa...', then the referent for the subject-Ø strategy controlling agreement on the perfective auxiliary ya could only have been 'the Tuareg'. This is because, in finite clauses at least,[14] Hausa does not permit the kind of "backward pronominalization" possible in such English sentences as "As soon as he$_j$ finished his$_j$ cigarette, Bond$_j$ leaned back...".

6. Summary

On the basis of the evidence presented in section 3, it seems reasonable to conclude that cognitive constraints relating to the encoder's ability to establish correct identification play an important part in determining selection of full NP's in preference to zero or pronominal anaphora. Thus, the occurrence, within the written narratives for example, of over 90% of all zero anaphors and almost 80% of pronouns following a discourse interval of no more than four clauses, compared with a corresponding figure of only 35% or so for coreferential NP's, is a distributional fact which would seem to be rooted in the capacity of human memory. Beyond this, however, I was able to demonstrate that the forces governing the pronoun:zero anaphora distribution in the direct object position are explicable not in terms of any universal cognitive constraints, but in terms of the salience:coding hypothesis already formulated to account for other language-specific facts -- prominent and persistent human referents are generally assigned more complex morphological coding than less salient nonhuman arguments.

Notes to Chapter 3

[1]Throughout this study I use terms such as "mention" and "control" to denote any coreferential antecedent device, whether zero anaphora, pronoun anaphora, or full NP.

[2]A regular feature of the written texts, but not the spoken variants, was the occurrence of direct speech quotations, and the clauses contained within tnese fragments have been included in the counts. This procedure is necessary because, as we shall see in detail in Chapter 5, the intervention of direct speech can exert some influence upon referential choice. The inclusion of individual direct speech clauses in the counts meant that the intervening clause scores for the written narratives are in general higher than the corresponding values for the oral narratives, as the following schema shows:

narrative clause	(a)	reference to X
direct speech clause	(b)	" "
narrative clause	(c)	reference to X

Inclusion of the direct speech clause (b) in the scores means that <u>two</u> clause boundaries separate the successive mentions of the referent X.

[3]The tally of 221 for pronoun anaphors in Table 3.1 includes, in addition to direct object pronouns, instances of indirect object, associative and possessive pronouns. Since, according to Clancy (1980:131), the only options used in the Japanese Pear Film accounts with any regularity were NP's and ellipsis, no figures are available for pronominal anaphora.

[4]Included in the direct object category are arguments which are formally expressed as possessive pronouns suffixed to certain classes of nonfinite verbonominals in the imperfective. Nonsubject zero anaphora is restricted primarily to the direct object of transitive verbs, and so I will restrict my present remarks to this phenomenon. It is worth noting, however, that the associative object of the so-called Hausa "causative" verb can be omitted in exactly the same kind of tight-knit environment which allows direct object zero-anaphora, though with nothing like the same text frequency. I encountered two examples -- (29c) in Chapter 2, and (4d) below:

(4) a. Wata rana wani bakauye ya labta wa
IS day-ADV IS villager-SUBJ he-PFV load to
b. jaka-n-sa itace,
donkeys-of-his-IO wood-DO
c. Ø ya nufo gari
SUBJ-Ø he-PFV head for town
d. don Ø ya sayar Ø.
so that SUBJ-Ø he-SUBJNCTV sell ASSOC-Ø

(Imam 1970:30)

'One day a (certain) villager loaded wood onto his donkeys and headed for town to sell Ø (it).'

The fully specified VP in (4d) would include the associative marker <u>da</u>, followed by an independent pronoun <u>shi</u>, i.e. <u>sayar</u> <u>da</u> <u>shi</u> (sell-CAUS with it-IND PRO-ASSOC) 'to sell it'.

[5]In order to provide a stronger numerical base for my claims, I scanned additional stories in Imam (1970) for further tokens of both direct object pronouns and zero anaphors. The total counts for these two forms in Figures 3.5-3.6 are, therefore, higher than those included in the preliminary scores given in Table 3.1.

[6]Wald (1979) describes a similar phenomenon in Mombasa Swahili, where there is a discernible correlation between the feature "human" and the occurrence of an "object marker". Corresponding scores from the oral narratives, though the sample is relatively small, are also in keeping with the claims, and are as follows:

DO Pro:	Human:	88.9%	(24/27)
	Inanimate:	11.1%	(3/27)
DO Zero:	Inanimate:	100%	(22/22)

[7]It would be useful to apply the same kind of test used by Li and Thompson (1979:322ff.) in their account of the zero:pronoun opposition in Mandarin Chinese -- remove all direct object pronouns from the texts and then ask native-speakers to indicate those locations where they felt pronouns were needed, i.e. in preference to zero anaphors. Such an experiment is beyond the scope of the present study, however.

[8]Both Paul Schachter and Russell Schuh (p.c.) have suggested to me that the verb <u>nema</u> 'look for' in (17a) might be intransitive in this context, in which case, of course, the question of this being a case of direct ob-

ject-Ø would not arise. Schachter, in fact, drew my attention to a tonally distinct (high-low tone) intransitive form of this verb listed by Abraham (1962:703) as meaning simply 'look'. However, both Hausa-speakers I consulted on this matter judged this to be the (low-high tone) transitive variant, adding that use of a (feminine singular) direct object pronoun ta, i.e. instead of zero, would have been perfectly acceptable.

[9]As Paul Schachter (p.c.) has correctly pointed out, since concord-drop and auxiliary-drop both differ from subject-Ø in that they do not entail the deletion of an argument, the question of potentially ambiguous arguments competing for reference is rendered immaterial for these two categories -- hence the absence of any values for this dimension in Table 3.3.

[10]In the future, e.g. (24d, g), the tense-apect marker and subject agreement morpheme, which together constitute the auxiliary, are transcribed as discrete elements in all but the first person singular and third person masculine singular.

[11]Newman and Schuh (1974:27-29), it is worth noting, hypothesize that although use of third-person subject-agreement on the imperfective auxiliary is a regular feature of modern Hausa, this represents an innovation, i.e., concord-drop was, historically, "concord non-insertion". Concord-drop is also possible in the largely-ignored "relative future" tense-aspect mentioned briefly in Abraham (1962:437). I encountered the following solitary example of the phenomenon in Imam (1970) -- buried within an idiom:

(28) a. "In Ø ta bi daga daga
if SUBJ-Ø she-PFV follow battle line
b. na ƙurya Ø-ka sha kashi."
of corner-SUBJ CONC-Ø-REL FUT suffer shit

(Imam 1970:9)

'When the frontline fighting is in progess, it's the people at the flanks who Ø will feel the brunt of it.'

[12]Although, as already noted, concord-drop and auxiliary-drop are not, unlike subject-Ø and direct object-Ø, analyzable as cases of "genuine" zero anaphora, I have taken the liberty of using the "Ø" symbol to denote these devices in both the citations and morphological glosses. This convention is merely a matter of notational convenience, and is not intended to indicate that they are of comparable status.

[13]Stylistic omission of the entire auxiliary is also possible with certain fixed, formulaic expressions in the subjunctive mood, e.g.,

(34) Allah Ø ji ƙa-n-sa.
Allah-SUBJ SUBJNCTV AUX-Ø feel mercy-of-him
'May Allah Ø have mercy on his soul.'

[14]"Backwards pronominalization" is acceptable, however, when a nonfinite (intransitive-motion) verbonominal is used, with a subjective possessive pronoun suffix, e.g.,

(38) a. Da zuwa-n-sa
with coming-of-his-VN-PARTICIPLE
b. sai Isa ya tuna da mafarki-n
then Isa-SUBJ he-PFV remember with dream of
jiya...
yesterday-ASSOC

'On his (Isa's) arrival, Isa remembered yesterday's dream.'

CHAPTER 4

CHARTING AND RETRIEVING REFERENTS WITH LEXICAL NP CONFIGURATIONS

1. Introduction

Once a referent has been introduced into the discourse, it is clear that a number of more, or less, explicit referential options are available for encoding subsequent mentions of the same entity, and in Chapter 3 we considered the distributional behavior of the two least overt categories -- zero anaphora and pronominal anaphoric reference. In this chapter we turn our attention to those more explicit lexical forms which can be exploited to specify an identifiable, definite referent. The combinational NP categories to be examined are, in order of presentation: (1) N(oun) + demonstrative (4.1.1); (2) N(oun) + determiner (4.1.2); (3) bare N(oun) (4.1.3); (4) proper names (4.1.4); (5) N(oun) + possessor (4.1.5). The kinds of questions we shall attempt to answer are: What communicative needs prompt an encoder to revert to full nominal mention in order to specify a referent? In particular, what are the rules which govern selection of one of the first three lexical configurations, i.e. a full nominal with or without an accompanying deictic morpheme?

Previous approaches to similar problems have, I believe, suffered from two analytical drawbacks. Either they have failed to consider the cross-text distribution of all the members of a given deictic system, and/or they have not taken into account a sufficiently wide range of factors responsible for the selection of different reference-types. Thus, the analysis presented in Jaggar (1983), while providing some insights into the factors conditioning selection of some of these referential items whose primary function is at the level of discourse, is nonetheless an oversimplification of what is a rich and subtle communicative sub-system.[1] Insufficient attention is paid to the highly explicit, heavy-coding N + demonstrative category, and, in keeping with the methodological guidelines set for the volume, the only motivating factors considered are the by now familiar ones of "time" and "interference" (cf. Chapter 3). Clancy's (1980) study of discourse strategies in the English and Japanese Pear Film narratives is a much more detailed investigation, but there is still no attempt to objectify the circumstances favoring selection of definite articles in preference to demonstratives -- the two categories are simply lumped together for the purposes of the analysis and treated as lexical NP's. Even Kirsner (1979:360) displays unusual imprecision when he writes that the choice between the two demonstratives of Modern Dutch "is like that between

either demonstrative and [the definite article] de, but with more nuance." Most uncharacteristically, Kirsner makes no attempt either to define the term "nuance" or invest the claim with any objectivity. This chapter seeks, therefore, to take up the desirable and challenging task of investigating a wider range of both coreferential NP-types and of factors motivating the discriminant choices encountered in natural discourse.

Section 2 provides the reader with some introductory material on the relative cross-text frequency of the various NP combinations. In section 3, I suggest a hypothesis to account for some of the variant distributional patterns, based upon the notion of the ease/difficulty of referent-identification, and proceed to validation of the claims in section 4.1. Finally, in section 4.2, I document cases of coreferential NP choices which are not explicable in terms of the proposed hypothesis -- the phenomenon of "new referent-anchoring" (4.2.1), and the coding of certain culturally salient referents (4.2.2).

2. The distributional data on NP categories: a quantitative overview

This section is designed to provide the reader with a brief, introductory profile of the formal distribution of the five NP configurations under consideration. Scrutiny

of the first five stories in Imam (1970:6-29) produced over 350 tokens in all; and the four oral narrative texts yielded a total of 71 such tokens. Table 4.1 summarizes tha cross-text data on the scatter of these NP categories.[2]

TABLE 4.1

Relative Frequencies of Coreferential NP Categories in Written and Spoken Narratives

NP-type	Written Narratives		Spoken Narratives	
	No.	%	No.	%
Simple N	139	38.7	30	42.3
N + determiner	48	13.4	22	31.0
N + demonstrative	49	13.6	4	5.6
N + possessor	39	10.9	11	15.5
Proper name	84	23.4	4	5.6
Totals	359		71	

This chapter is largely concerned with the contrasting behavioral characteristics of the first three categories listed in Table 4.1, i.e. simple noun, noun + determiner, and noun + demonstrative, and the statistics given in

Table 4.1 permit several immediate observations. Firstly, in both the written and spoken narrative texts, use of a simple nominal accounted for the majority of coreferential NP choices -- a markedly similar 38.7% and 42.3% respectively. Secondly, whereas demonstratives and determiners occurred with almost identical frequency in the written narrative sample -- 13.6% and 13.4% respectively of all NP mentions -- determiners had a noticeably higher percentage occurrence in the oral narratives -- 31.8% compared with 5.6% for demonstratives. These and other related matters are discussed in more detail in the relevant sections below.

3. The hypothesis

In order to provide a natural basis for explaining at least some of the discriminant NP choices encountered in the corpora, I have adapted claims advanced in García (1975:65) and Kirsner (1979:358). The working hypothesis states that the meanings signaled by the three NP categories are organized in terms of a scale of "DEIXIS", operationally defined as "the force with which the encoder points out a particular referent, thereby both urging and assisting the decoder to find and correctly identify the same referent". At the two extremes of this graded continuum, we find the noun + demonstrative configuration

denoting STRONG or HIGH DEIXIS, the category of bare nominal signaling WEAK or LOW DEIXIS, and the noun + determiner grouping occupying a somewhat intermediate position with respect to its deictic strength, let us say it maps INTERMEDIATE DEIXIS. This arrangement is represented in the schema given in (1):

(1) STRONG DEIXIS ---------- WEAK DEIXIS

N + demonstrative N + determiner Simple N

It now becomes necessary to ask precisely why an encoder might wish to convey stronger deictic signals. I suggest one general discourse-based circumstance -- allowing, as ever, for the possibility that more no doubt exist[3] -- in which the encoder might wish to alert a decoder to a particular referent, prompting the decoder and helping in the cognitive task of referent-identification. This is:

(2) EASE/DIFFICULTY OF REFERENT-IDENTIFICATION (cf. section 4.1). The referent has been absent from the discourse for some time and so may no longer be presumed to be fully activated in the decoder's "consciousness", (Chafe 1974, 1980a; Bernardo 1980), and/or there may be interference from maximally

similar, potentially ambiguous referents. These two parameters -- "time" and "interference" -- combine to render the decoder's task of referent-identification more difficult. The hypothesis will therefore predict, following Givón (1983), a tendency for the encoder to use the heavier coding associated with HIGH DEIXIS if he/she considers the decoder might otherwise have difficulty uniquely accessing the intended referent.

4. Empirical verification of the claims

We now proceed to validation of the above deixis:coding hypothesis via a detailed assessment of the relation between the factor of relative ease/difficulty of referent-identification and the five NP-configurations under consideration, concentrating mainly on the categories N + demonstrative, N + determiner, and bare nominal. Comparison of these configurations is of particular interest for they all involve decisions to mark, or not in the simple N case, a lexical NP mention with one of the two deictic operators available, i.e. demonstrative or determiner. The category of proper name, on the other hand, rarely permits such marking; and the class of N + possessor constitutes, as we shall see, a rather special and narrowly-defined referential category.

4.1. Ease/difficulty of referent-accessibility and morphological coding

As already noted (cf. Chapter 3 for references), a number of linguistic and psycholinguistic studies have independently shown that the passage of time and/or the intervention of potentially ambiguous referents have a significant influence upon the decoder's ability to process incoming referential information. It was also demonstrated in Chapter 3 that the general class of coreferential NP's was characterized by significantly higher averages for both intervening clause boundaries and potential confusers than was the case with the attenuate pronoun and zero anaphoric forms. Looking within this generalized NP domain, we might anticipate a correlation between the choice of a particular coreferential strategy and clause distance to prior mention and/or the number of intervening potential confusers, with the high-density coding device of N + demonstrative correlating with higher quotients for the above two measures, followed by the category N + determiner, and finally the strategy of bare nominal.

Table 4.2 now furnishes data on all five categories with respect to the selected dimensions of time, i.e. the number of clause boundaries between two successive mentions of a given referent, and interference, i.e., the number of maximally similar confusers present between two

mentions -- where the second of the two mentions is selected from one of the NP configurations under consideration. In Table 4.2, the values above the line represent the averages computed for each of the two measures, and the numbers below the line indicate the type-token counts. Note that the totals for the categories of N + determiner and N + demonstrative are higher than those given in Table 4.1 -- this is because additional texts were searched in order to collect token-counts comparable to those recorded for bare nouns.

TABLE 4.2

Distribution of NP Categories with respect to Time and Interference

NP Type	Written Narratives		Spoken Narratives	
	Av. # of clauses	Av. # of confusers	Av. # of clauses	Av. # of confusers
Noun	11.3 / 139	0.7 / 139	6.7 / 30	0.6 / 30
N + DET	13.9 / 107	0.7 / 107	7.9 / 22	0.5 / 22
N + DEM	15.2 / 112	0.9 / 112	14.7 / 4	0.8 / 4
N + POSS	22.9 / 39	0.4 / 39	10.2 / 11	0.9 / 11
Proper name	11.1 / 84	0.6 / 84	8.0 / 4	0.3 / 4

Several noteworthy patterns emerge from the raw figures in Table 4.2:

(3) With respect to the dimension of time, one of the major communicative motivations for use of the relatively heavyweight strategy of N + demonstrative -- the coding of a discourse referent which is at some distance from its immediately preceding control -- is

reflected in the fact that, in the written and spoken narratives respectively, averages of 15.2 and 14.7 clauses intervened between successive mentions, the highest of any of the three main NP categories. It is worth noting too that these averages would have been higher had it not been for the effect of the strategy of "new referent-anchoring" discussed in section 4.2.1. At one remove lower on the scale we find the category N + determiner with slightly lower averages of 13.9 and 7.9 clauses respectively, with the bare nominal category showing the lowest score of the three for intervening clauses -- 11.3 (written) and 6.7 (oral).[4] In general, the average confuser rates do not display any significant variation.

Tables 4.3-4.6 now provide a more detailed breakdown of the distribution of the five NP categories with regard to time and interference, and together with Table 4.2 provide the quantitative base upon which the remainder of the discussion is presented.

TABLE 4.3

Distribution in Written Narratives of NP Categories with respect to the Number of Clauses present between Two Successive Mentions of a Referent.

NP-type	Number of intervening clauses							
	1-4		5-10		11-20		21+	
Bare noun	53 / 139	38.1%	52 / 139	37.4%	22 / 139	15.9%	12 / 139	8.6%
N + DET	36 / 107	33.7%	35 / 107	32.7%	21 / 107	19.5%	15 / 107	14.1%
N + DEM	35 / 112	31.3%	30 / 112	26.8%	25 / 112	22.3%	22 / 112	19.6%
N + POSS	11 / 39	28.2%	13 / 39	33.3%	8 / 39	20.5%	7 / 39	18.0%
Proper name	34 / 84	40.5%	28 / 84	33.3%	15 / 84	17.9%	7 / 84	8.3%

TABLE 4.4

Distribution within Spoken Narratives of NP Categories with respect to the Number of Clauses present between Two Successive Mentions of a Referent.

NP-type	Number of intervening clauses							
	1-4		5-10		11-20		21+	
Bare noun	16/30	53.3%	9/30	30.0%	3/30	10.0%	2/30	6.7%
N + DET	13/22	59.1%	4/22	18.2%	5/22	22.7%		
N + DEM	2/4	50.0%					2/4	50.0%
N + POSS	5/11	45.4%	4/11	36.4%			2/11	18.2%
Proper name			3/4	75.0%	1/4	25.0%		

A number of points emerge from the statistics given in Tables 4.3 and 4.4. Briefly, we may note, first of all, that, with the marginal exception of the N + possessor category, the largest number of tokens in all five NP categories in both the written and oral texts tend to occur after 1-4 clauses have elapsed since prior discourse mention of a referent, a fact which is basically congruent with Clancy's (1980:137ff.) findings for English.[5] Ob-

serve too, within the written narratives (Table 4.3), that the high deixis-coding demonstrative has the highest average counts falling within the two largest clause distance groups -- 22.3% and 19.6% in the (11-20) and (21+) clause range -- followed by the intermediate deixis N + determiner category -- 19.5% and 14.1% -- and finally the low deixis-signaling simple N category -- 15.9% and 8.6%. And in the oral texts (Table 4.4), we encounter an unusually high percentage of N + determiner tokens occurring after a 1-4 clause gap -- 59.1%, compared with 33.7% in the corresponding written narratives.

TABLE 4.5

Distribution within Written Narratives of NP Categories with respect to the Number of Potential Confusers Intervening between Two Successive Mentions of a Referent.

NP-type	Number of intervening confusers					
	0		1		2-4	
Bare noun	67 / 139	48.2%	53 / 139	38.1%	19 / 139	13.7%
N + DET	61 / 107	57.0%	30 / 107	28.0%	16 / 107	15.0%
N + DEM	52 / 112	46.4%	39 / 112	34.8%	21 / 112	18.8%
N + POSS	30 / 39	76.9%	7 / 39	18.0%	2 / 39	5.1%
Proper name	43 / 84	51.2%	27 / 84	32.1%	14 / 84	16.7%

TABLE 4.6

Distribution within Spoken Narratives of NP Categories with respect to the Number of Potential Confusers Intervening between Two Successive Mentions of a Referent.

NP-type	Number of intervening confusers					
	0		1		2-4	
Bare noun	20 / 30	66.7%	8 / 30	26.7%	2 / 30	6.6%
N + DET	17 / 22	77.3%	4 / 22	18.2%	1 / 22	4.5%
N + DEM	2 / 4	50.0%			2 / 4	50.0%
N + POSS	4 / 11	36.4%	6 / 11	54.5%	1 / 11	9.1%
Proper name	3 / 4	75.0%	1 / 4	25.0%		

The data in Tables 4.5 and 4.6 permit the following generalizations:

(4) For all the coreferential NP configurations in the written narratives, and most of those in the spoken narratives, the largest number occurred without the intervention of any maximally similar and so potenti-

ally confusing argument.

(5) Although the numerical differences are not too significant, within the written texts (Table 4.5), the high deixis category of N + demonstrative has the largest tabulated percentage in the maximum 2-4 confuser group -- 18.8% -- followed by N + determiner (15.0%) and bare N (13.7%).

Since the formal distribution of nominal coreference is apparently sensitive to the influence of time and, though to a lesser extent, interference, I shall now exemplify and discuss, category by category, those NP choices whose motivation seems to be related to cognitive constraints upon the decoder's presumed ability to decipher reference. In some cases, we shall also note certain semantic facts which contribute to the options selected.

4.1.1. Noun + demonstrative

Table 4.7 lists the so-called set of "nan-demonstratives" which are here considered.[6]

TABLE 4.7

The nan-marked Demonstratives

	Prenominal		Postnominal
Masc. sing.	wànnan	NP	NP-ǹ nan
Fem. sing.	wànnan	NP	NP-ɾ̀ nan
Plural	wàdànnan	NP	NP-ǹ nan

Hausa demonstratives occur in immediate prenominal or postnominal position. When preposed, it is the full form of the demonstrative which occurs -- the form also used as an anaphoric pronoun.[7] When postposed, an attenuate form is used, and the nominal carries a gender/number-sensitive suffix which is the probable source of the determiner element discussed in section 4.1.2 below.[8]

The nan-demonstratives in Table 4.7 may be exploited to code referents which are either visible or nonvisible-referential, and it is this latter temporal-referential function which is predictably encountered in narrative contexts of the kind investigated here.[9] When specifying entities locatable in physical space, it is of interest to note that these demonstratives encode referents which are closer to the addressee than to the speaker -- a usage not reported in descriptive grammars of the language[10] -- and it may be that it was this same addressee-oriented loca-

tive function which led to this form taking on a temporal-referential, but still essentially addressee-centered role, i.e. a common semantic function of "closeness/tangibility" is being exploited.

Inspection of the written narrative counts in Table 4.2 reveals that, of the three categories, the N + demonstrative configuration has the highest average scores for both the number of intervening clauses -- 15.2 -- and the number of intervening referents -- 0.9. Tables 4.3 and 4.4 provide more detailed summaries of these distributional data. Table 4.3 shows that this same category displays the largest incidence of tokens within both the 11-20 clause grouping (22.3%) and the maximal 21+ clause set -- 19.6%. And Table 4.5 reveals that the category N + demonstrative also has the highest percentage of potential confusers in the 2-4 range -- 18.8%. The token count of four for the corresponding category in the oral texts was really insufficient to permit any useful comparisons, though it is perhaps worth noting in passing that it did display the highest average for intervening clauses (cf. Table 4.4) -- 14.7 -- and 50% of the type-tokens counted fell into the 21+ clause range. These distributions suggest, I believe, that one of the major factors motivating selection of a coreferential high deixis demonstrative is a relatively substantial degree of discontinuity within the text. Fragment (6) below illustrates the kind of

disruptive environment in which the choice of a demonstrative is arguably attributable to the degree of discontinuity involved:

(6) a. ø Ya-na ajiye-wa ø,
SUBJ-ø he-IMPFV put down-VN DO-ø
b. ø ya ba da baya,
SUBJ-ø he-PFV give back
c. sai Lamiri ya ce wa
then Lamiri-SUBJ he-PFV say to
dan ƙane-n nan nasa...
DIMIN-of younger brother-DET DEM of his-IO

(Imam 1970:32)

'He (the servant) was putting (the water) down, and had withdrawn, then Lamiri said to that little younger brother of his...'

In (6c), the demonstrative-equipped indirect object NP dan ƙanen nan 'that younger brother' has a look-back of 30 clauses or so to its prior discourse control,[11] and there are also several potential interferers lurking in the vicinity, including the maximally-similar subject 'the servant' referred to in (6a-b). Hence, I would claim, the decision to use the escalated demonstrative device. Consider also extract (7), this time illustrating an animal referent with accompanying demonstrative:

(7) a. Yaro ya kwashe labari-n dodanniya-n nan
boy-SUBJ he-PFV collect news-of spirit-DET DEM
b. da yadda ø ta-ke rikida
and how SUBJ-ø she-IMPFV change form-VN
c. ø ya gaya wa Jimrau,
SUBJ-ø he-PFV tell to Jimrau-IO
d. ø ya kuma gaya masa
SUBJ-ø he-PFV and tell to him-IO PRO

dawa-n da Ø ta-ke,
bush-DET REL SUBJ-Ø she-IMPFV

e. Ø ya ba shi dabaru-n da Ø
SUBJ-Ø he-PFV give him-DO PRO plans-DET REL SUBJ-Ø
zai yi duka,
he-FUT do all

f. **kyanwa-n nan** ta taimake shi...
cat-DET DEM-SUBJ she-SUBJNCTV help him-DO PRO

(Imam 1970:50)

'The boy told Jimrau the story of that evil spirit and how she could change shape, and he told him the remote area where she was, and he gave him all the plans he should use, so that that cat might help him...'

The demonstrative-marked animal subject NP **kyanwan nan** 'that cat' in (7f) has been absent for 11 clauses -- a discourse interval which also includes an episode-shift and a block of quoted direct speech (cf. Chapter 5).[12] It is also possible that the presence of another feminine singular participant -- dodanniyan nan 'that (female) evil spirit' referred to in (7a, b, d) -- also contributes to the decision to use an unambiguous N + demonstrative strategy in (7f).

Finally, extract (8) contains an example of a demonstrative-coded inanimate:

(8) a. Ɗa-m barga ya wuce
son-of stable-SUBJ he-PFV pass by

b. Ø bai ko amsa mata ba.
SUBJ-Ø NEG-he-PFV even answer to her-IO PRO NEG

c. Ø Ya je
SUBJ-Ø he-PFV go

d. Ø ya tuɓe **kaya-n** nan
SUBJ-Ø he-PFV take off **clothes-DET** DEM-DO

e. Ø ya ɓoye Ø...
SUBJ-Ø he-PFV hide DO-Ø

(Imam 1970:44)

'The stable boy passed by and didn't even answer her. And he went and took off those clothes and hid (them)...'

In clause (8d), the direct object N + demonstrative **kayan nan** 'those clothes' resurfaces after being "off-stage" for some 45 clauses.

4.1.2. Noun + determiner

As already noted, historically the determiner morpheme is probably a morphologically truncated form of the N + demonstrative construction described above in section 4.1.1. A nominal marked with a determiner, therefore, carries a gender/number-sensitive suffix, but without the postposed demonstrative element, most probably the high-tone, referential remote-distal can variant discussed briefly in fn. 9 above.[13] The determiner parallels the demonstrative in that it is regularly used in discontinuous contexts, i.e. contexts in which the passage of time, as measured in terms of intervening clauses, and/or the presence of maximally similar referents, persuade the encoder that these particular strategies are more appropriate if the decoder is to be able to assign unique and proper identity to the referent so coded. Unlike the demonstrative, however, the determiner cannot be exploited to encode distal/proximal distinctions in time or space.

The cross-category data in Table 4.2 suggest some support for the claim that the determiner is manipulated to signal what has been termed intermediate deixis -- it has a mean of 13.9 intervening clauses, i.e. falling between the corresponding scores for demonstrative-coded nouns (15.2), and bare nouns (11.3). The breakdowns in Table 4.3 provide more detailed evidence -- type-token percentages of 33.7% and 32.7% respectively for (1-4) and (5-10) intervening clauses, compared with 31.3%/26.8% for the corresponding N + demonstrative category and 38.1%/37.4% for simple N's. And at the other end of the scale, the category N + determiner displays percentages of 19.5%/14.1% in the (11-20) and (21+) clause range, compared with 22.3%/19.6% and 15.9%/8.6% for the categories N + demonstrative and bare noun respectively. And looking at the compilations for the oral narratives, Table 4.2 also shows that, in general, occurrences of the N + determiner category are found over marginally longer discourse stretches than are bare N's -- a mean of 7.9 clauses compared with 6.7, though the difference is certainly not significant. Notice, however, the jump to 14.7 clauses recorded for the demonstrative category.

Consider excerpt (9):

(9) a. Sarki ya tashi,
emir-SUBJ he-PFV get up
b. Ø ya shiga gida duk rai a ɓace,
SUBJ-Ø he-PFV enter house all mind at spoil-STAT

c. Ø ya kira 'ya-r da
SUBJ-Ø he-PFV summon daughter-DET-DO and
uwa-r-ta...
mother-of-her-DO

(Imam 1970:24)

'The emir got up, entered the house most distressed and called the daughter and her mother...'

The direct object N + determiner example in question -- 'yar 'the daughter' in (9c) -- has a hefty look-back of more than 150 distinct clauses to its previous appearance in the story.[14]

Fragment (10) contains an example of the way in which an intervening potential confuser can contribute to use of a determiner:

(10) a. Ø Ya kira wani ɗa-n
SUBJ-Ø he-PFV summon IS DIMIN-of
ƙane-n-sa ƙarami,
younger brother-of-his small-DO
b. Ø suka shiga mota-r tajiri-n...
SUBJ-Ø they-PFV enter car-of merchant-DET

(Imam 1970:32)

'He (Lamiri) called one of his younger brothers and they got into the merchant's car...'

Although the possessor N + determiner tajirin 'the merchant's' in (10b) reactivates a mention only 5 clauses earlier, the presence of the maximally similar referent ƙanensa 'his younger brother' in (10a) goes some way to explaining selection of the escalated strategy in order to avoid potential ambiguity as to ownership of the car. And

Russell Schuh (p.c.) has also made the point that putative use of a simple, non-modified nominal in this context, i.e. motar tajiri, could permit a generic-sortal interpretation, denoting the type of car as opposed to the specific possession of a specific referent. Semantic input of this nature clearly needs to be considered in the search for factors determining ongoing referential choice -- cf. my remarks on the marking of kin-terms in fn. 11.

Fragment (11) illustrates a determiner-marked inanimate subject:

(11) a. Ko da Ø ta cinye dodanniya,
As soon as SUBJ-Ø she-PFV eat up spirit-DO
b. sai Ø ya ji
then SUBJ-Ø he-PFV hear
c. gida-n ya ɗauka "....."
house-DET-SUBJ it-PFV take

(Imam 1970:51)

'As soon as she (the cat) had eaten up the evil spirit, he (Jimrau) heard the house say ".....".'

Prior mention of the determiner-equipped NP gidan 'the house' in (11c) is over 70 clauses to the left.

Finally in this section, recall that the average look-back for determiner-coded referents in the oral texts was an unusually low 7.9 clauses -- much closer, in fact, to the 6.7 computed for bare nouns than the 14.7 tally for demonstratives (cf. Table 4.2). I believe an explanation is available, and it relates to the discernible tendency amongst young, bilingual Hausa-English speakers acquainted

with the use of the English definite article to mark identifiable referents with a determiner, i.e. in preference to using a simple noun.[15] Extract (12) illustrates:

(12) a. To, ashe Ø ya bar hula-r-sa
OK really SUBJ-Ø he-PFV leave cap-of-his-DO
a wuri-n da aka yi karo-n,
at place-DET REL IMPERS-PFV do collision-DET
b. sai wani yaro ya ga hula-r.
then IS boy-SUBJ he-PFV see cap-DET-DO

(Speaker 1)

'OK, he (the boy on the bike) had left his cap where the collision had happened, then a boy saw the cap.'

This fragment contains two determiner-marked inanimates -- karon 'the collision' in (12a), and hular 'the cap' in (12b), and although the first item has a reasonably lengthy look-back of 12 clauses to prior control, the (12b) hular reference follows hard on the heels of its previous mention in (12a).

4.1.3. Simple nominal

A glance at Table 4.2 shows that, of the three configurations, the low-deixis category of bare N displays the smallest averages for intervening clauses -- 11.3 and 6.7 in the written and spoken texts repectively. And Table 4.3 shows that the percentage of simple N's occurring within the ranges (11-20) and (21+) clauses in the written narratives is also the lowest recorded -- 15.9% and 8.6%.

With respect to the dimension of intervening confusers, Table 4.2 shows that the simple N category is closer to the N + determiner group -- averages of 0.7 (written) and 0.6 (spoken), compared with 0.7/0.5 for determiner-marked nouns -- than it is to the demonstratives (0.9/0.8) -- though the numerical differences are not too striking. Fragments (13-15) provide exemplification:

(13) a. ∅ Ya sa
SUBJ-∅ he-PFV cause
b. sarki-n dogarai ya bi
chief-of bodyguards-SUBJ he-SUBJCTV follow
shi
him-DO PRO
c. inda ∅ ya binne ∅,
where SUBJ-∅ he-PFV bury DO-∅
d. ∅ ya tono ∅
SUBJ-∅ he-SUBJCTV dig up DO-∅
e. ∅ ya kawo ∅.
SUBJ-∅ he-SUBJCTV bring back DO-∅
f. Aka ba mai kudî abi-n-sa.
IMPERS-PFV give one with money-IO thing-of-his-DO

(Imam 1970:29)

'He (the judge) had the chief bodyguard follow him (the thief) to where he had hidden (the money) for him to dig (it) up and bring (it) back. **The original owner** was then given his money.'

In clause (13f), the indirect object noun **mai kudî** 'money owner' picks up a previous reference 30 clauses to the left -- an N + possessor pronoun <u>bakonsa</u> 'his (the chief butcher's) guest' -- and four potential confusers intervene between the two successive mentions.

Fragment (14) is taken from one of the oral narratives:

(14) a. Da muka jawo ta,
When we-PFV pull it-DO PRO
b. muka canja **taya...**
we-PFV change **tire-DO**

(Speaker 1-2)

'When we pulled it (the car) out, we changed **the tire...**'

In (14b), the direct object inanimate noun taya 'tire' has a look-back of 60 separate clauses. And in (15c):

(15) a. Da gani-n-sa
with seeing-of-him-PART
b. sai Isa ya tuna da
then Isa-SUBJ he-PFV remember with
mafarki-n jiya,
dream-of yesterday-ASSOC
c. ∅ ya ce wa **wanzami...**
SUBJ-∅ he-PFV say to **barber-IO**

(Imam 1970:12)

'On seeing him (the Tuareg), Isa remembered yesterday's dream and said to **the barber...**'

the indirect object N **wanzami** 'the barber' has a look-back of 12 clauses to its prior control, and this discourse gap includes mention of the potentially ambiguous argument <u>Buzu</u> 'the Tuareg' referred to in (15a).

4.1.4. Proper name

This category has not been grouped with the three NP combinations discussed above because none of the tokens noted in the corpora were coded with either of the two deictic morphemes.[16] Intuitively, one might expect proper

names to behave in a fashion comparable to the class of simple N's, and indeed the written narrative data in Table 4.2 reveal that they display clause averages which are closest to those of bare N's -- an average of 11.1 intervening clauses, the lowest, in fact, of all five NP categories. Consider, therefore, extract (16):

(16) a. Sarki ya yi murmushi,
emir-SUBJ he-PFV do smile
b. Ø ya dubi **Nahi'u Hankali**
SUBJ-Ø he-PFV look at **Nahi'u Hankali-DO**
c. Ø ya ce...
SUBJ-Ø he-PFV say

(Imam 1970:9)

'The emir smiled, looked at **"Nahi'u-the-Careful"** and said...'

The direct object proper name **Nahi'u Hankali** 'Nahi'u-the-Careful' in (16b) is previously referred to some 63 clauses to the left, and there are four maximally similar arguments intervening between the two mentions, including Nahi'u's two brothers, all of which could potentially have answered to a less overt reference-type, e.g. pronoun, had it been used. This type of interference thus seems sufficient to elicit use of the full proper name, for the putative selection of a less overt referential form would have yielded genuine ambiguity of reference.[17]

4.1.5. Noun + possessor

Table 4.2 shows that nouns qualified by a following possessor argument, whether a full nominal or pronominal suffix, display the highest averages of any lexical NP category -- 22.9 and 10.2 intervening clauses for the written and spoken texts, and an average of just less than one potential confuser in the spoken stories. A possible explanation of the relatively high clause values is, I believe, available. The vast majority of possessed N's counted, excepting a few kin-terms, were inanimate physical objects occupying nonsubject roles, and since such entities tend, unlike "stage-holding" human participants, to appear only as intermittent "props" in the narrative flow, the discourse gaps between consecutive mentions tend to be rather substantial.[18] Extracts (17-18) contain typical examples:

(17) a. Sai Kalala ya runtuma a guje
then Kalala-SUBJ he-PFV run at flee-STATIVE
da **wuka-r-sa** a hannu...
with **knife-of-his-ASSOC** in hand-LOC

(Imam 1970:19)

'Then Kalala ran off at full pelt with **his knife** in his hand...'

The associative-possessed NP **wukarsa** 'his knife' in (17a) has been absent from the plot for more than 90 distinct clauses. And in (18d):

(18) a. ...Ø ya tabbata
SUBJ-Ø he-PFV be sure
b. gobe alƙali ɗaure shi
tomorrow judge-SUBJ imprison-VN-TOPIC him DO PRO
zai yi
he-FUT do
c. in an sami
if IMPERS-PFV find
d. sanda-n-sa ya ƙaru da rabi-n
stick-of-his-SUBJ it-PFV increase with half-of
taƙa.
pace-ASSOC

(Imam 1970:28)

'...he (the thief) was sure that the judge would imprison him the following day if it was found that **his stick** had increased in length by half a pace.'

the possessed NP **sandansa** 'his stick' has a look-back of some 20 clauses.

4.2. Coreferential NP choices not determined by text discontinuity

The discussion so far has been limited mainly to illustrating, in a sometimes rather obvious fashion, the kinds of disruptive contexts which typically induce selection of a full coreferential nominal, with or without an accessory deictic. The analysis I formulated in Jaggar (1983), moreover, in keeping with the general methodological bias of the volume, dealt almost exclusively with the cognitive factors of time and/or interference, and I was concerned at the time about this unidimensional view of what is in reality a much more sophisticated referential domain. It has since become abundantly clear that this

approach to the data was both too narrow and too mechanical, failing to take account of other important determinants of referential choice; in other words, the dimensions of time and interference cannot be used to explain all the NP referential types encountered in Hausa narrative structure, and indeed we have already noted cases where semantic facts appear to be of some relevance to these decisions. A more valid investigation, therefore, will necessarily have to consider those discourse contexts in which the time variable in particular -- as measured in terms of the number of intervening clauses -- cannot be held responsible for the ongoing referential choices made.

There appear to be three major factors precipitating selection of NP coreference over unusually short intervals of time, and relating to: (1) Contexts in which there is an interchange of subject arguments ("subject-switching") and/or where some kind of episode-boundary has intervened; (2) Contexts in which the encoder feels a need to firmly establish a recently-introduced referent in the mind of the decoder and accomplishes this communicative task by using an escalated NP strategy (**"new referent-anchoring"**); (3) The more marginal tendency to encode culturally-prominent referents with heavy morphology. The latter two phenomena, it may be noted, both represent further violations of Givón's (1983) hypothesis -- stated and discussed in Chapter 3 -- which claims that formal coding devices

covary in their size and complexity with the continuity/predictability of the referent. Discussion of the first phenomenon -- "subject-switching" and NP coreference -- is left for the next, and final chapter, which also looks at the impact of stronger discourse breaks, e.g. episode boundaries, on referential choice. The present chapter now concludes with some observations on the matter of "new referent-anchoring" (4.2.1), and also on the coreferential coding of culturally-salient humans (4.2.2).

4.2.1. New referent-anchoring

By **"new referent-anchoring"** I mean the use of an accompanying deictic -- most typically a demonstrative in the sample texts investigated -- to spotlight a newly-introduced referent and so anchor it as a salient entity in the consciousness of the decoder, usually because that same argument is being prepared for active deployment/further participation in the ensuing discourse. This anchoring takes place early on in the discourse career of a newly-introduced referent, following hard on the heels of first lexical mention, and it is a function which is in some respects, therefore, analogous to that of the "Indefinite Specifier" (cf. Chapter 2) in the sense that it raises the decoder's expectation that more information is about to be added about the referent so marked. I record-

ed 13 recognizable cases of new-referent anchoring in the first 11 stories of Imam (1970:6-65), 9 of which were "flagged" with a demonstrative, and the remaining 4 with a determiner. Of the 9 demonstrative-marked items, there were 3 human subject referents, 2 human nonsubject arguments, and 4 nonhuman nonsubject cases; and of the 4 determiner-coded anchoring tokens, 3 involved human referents -- 2 subjects, and 1 nonsubject -- with 1 direct object inanimate.[19] Fragment (19) contains two examples of newly-introduced human referents, each of which is "moored" on second lexical mention with an accompanying demonstrative:

(19) a. A-na nan, ran nan wani
IMPERS-IMPFV there-LOC day DEM IS
makwabci-n-sa ya aza wa
neighbor-of-his-SUBJ he-PFV load to
jaka-n-sa ice,
donkeys-of-his-IO wood-DO
b. Ø ya nufi gari.
SUBJ-Ø he-PFV head for town
c. Ø Ya-na isa,
SUBJ-Ø he-IMPFV approach-VN
d. sai Ø ya tarar
then SUBJ-Ø he-PFV find
e. a-na ta jaje
IMPERS-IMPFV continue sympathy
f. an yi wa wani attajiri sata-r
IMPERS-PFV do to IS merchant-IO stealing-of
jaka uku.
bag three
g. Da **mutumi-n** nan ya ji haka,
when **man-DET** **DEM-SUBJ** he-PFV hear this
h. wai shi ba'a sai Ø ya tafi
that he-IND PRO mockery then SUBJ-Ø he-PFV go
wuri-n **attajiri-n** **nan,**
place-of **merchant-DET** **DEM**
i. Ø ya ce...
SUBJ-Ø he-PFV say

(Imam 1970:31)

'Well there they were, one day[20] a neighbor of his loaded wood on his donkeys and made for town. As he was approaching he discovered that sympathy was being expressed for a merchant who'd had three bags of money stolen. When **this man** heard this, bent on mockery he went to **that merchant** and said...'

The first demonstrative-anchored new referent -- the human subject NP **mutumin nan** 'this man' in (19g) -- is immediately reactivated despite only a temporary absence since prior mention in (19d) and is deployed in the subject slot until the end of the paragraph-marked episode. The second case -- the demonstrative-anchored possessor NP **attajirin nan** 'that merchant' in (19h) -- also looks back a mere 3 clauses to its first lexical mention in the story, and also embarks upon a pivotal career, in this case until the end of the story. In both the above cases, I would contend that it is the communicative need to anchor a salient, deployable referent in the decoder's mind which explains the occurrence of the accessory demonstratives.[21] It is precisely these types of examples, moreover, which are problematical for Givón's (1983) continuity:coding hypothesis -- the discourse absence of the referents in (19g, h) is negligible, and there are no maximally similar intervening arguments competing for reference -- yet they are equipped on second lexical mention with a high-density demonstrative. It may be noted too that the averages computed for the demonstrative cate-

gory -- 15.2 clauses intervening between successive mentions of a referent (cf. Table 4.2) -- would have been higher were it not for the effects of this anchoring function.

Consider too excerpt (20):

(20) a. To, akwai wata dodanniya wadda ta-ke
OK, EXIST IS ogress REL she-IMPFV
tare fatake,
intercept traders-DO
b. Ø ta-na kashe su
SUBJ-Ø she-IMPFV kill them-DO-PRO
c. Ø ta cinye Ø
SUBJ-Ø she-SUBJCTV eat up DO-Ø
d. Ø ta ƙwace kaya-n,
SUBJ-Ø she-SUBJCTV steal loads-DET-DO
e. Ø ta tara dukiya mai yawa
SUBJ-Ø she-SUBJCTV amass wealth with plenty-DO
da kudi jakunkuna a gida-n jibge.
and money bags-DO in house-DET pile up-STAT
f. Yaro ya kwashe labari-n **dodanniya-n nan...**
boy-SUBJ he-PFV tell news-of **ogress-DET DEM**

(Imam 1970:50)

'OK, there was a certain ogress who was intercepting traders, and she would kill them, eat them up, steal the loads and amass great wealth and bags of money piled up in the house. The boy related the story about **this ogress...**'

'The ogress' is first introduced into the story with the IS-marked NP wata dodanniya in (20a), and persists in the subject role until clause (20e). The second lexical mention, only one clause later in (20f), is again equipped with a full demonstrative -- the possessor N + demonstrative **dodanniyan nan** 'this ogress'. Although a switch of subject has taken place in this clause, it remains the

case that selection of a less overt anaphoric strategy, e.g. a feminine gender possessive pronoun suffix, would not have led to ambiguity in any sense,[22] and again I would contend that the demonstrative being is manipulated to anchor the freshly-introduced argument in the addressee's mind, ready for eventual, and often immediate, discourse deployment.

Examples (21) and (22) show that nonhuman (animal or inanimate) referents may also be anchored in the same fashion in order to prepare them for deployment as important arguments within the discourse:

(21) a. ...amma duk da haka a ciki-n
but in spite of this at inside-of
balbelu-n nan
egrets-DET DEM
b. akwai wata wadda sarki ya-ke so.
there is IS REL emir he-IMPFV like-VN

(Imam 1970:7)

'...but in spite of this, amongst **these cattle egrets** there was one that the emir liked.'

In (21a) the demonstrative-anchored prepositional NP **balbelun nan** 'these cattle egrets' has a look-back of only 3 clauses to its first discourse mention, and 'the cattle egrets' themselves represent an important backdrop to the unfolding plot, with one of them personified and taking on various human attributes, including the power of speech.

Consider also (22):

(22) a. Ashe guda Ø-na da kudi,
strangely one-SUBJ CONC-Ø-IMPFV with money-ASSOC
b. Ø bai gaya wa kowa ba
SUBJ-Ø NEG-he-PFV tell to anyone-IO NEG
ciki-n 'yanuwa-n nan.
inside-of relatives-DET DEM
c. Sai can da dare Ø su-na hira,
then later at night-ADV SUBJ-Ø they-IMPFV chatting
d. Ø ya ɗauko **kudi-n nan**
SUBJ-Ø he-PFV take out **money-DET DEM-DO**
e. Ø ya ƙidaya su sule goma
SUBJ-Ø he-PFV count them-DO PRO shillings ten
sha biyar...
and five

(Imam 1970:27)

'Strangely enough one of them had some money, but hadn't told any of those brothers. Then later on in the night they were chatting, and he took out **this money** and counted out fifteen shillings...'

Once again, there is only a very short interval between the demonstrative-anchored argument -- in this case the direct object inanimate NP **kudin nan** 'this money' in (22d) -- and its immediately preceding first mention in (22a). Nor is there any possibility of real ambiguity arising had a reduced reference-type been used, e.g. a pronoun or zero anaphor. Again too, the now firmly-rooted new referent, though inanimate, proceeds to play an important communicative role, providing background motivation for a number of twists and turns in the unfolding narrative.

Fragment (23) exemplifies a similar case from one of the Pear Film accounts:

(23) a. Baya-n Ø ya hau itace,
after-of SUBJ-Ø he-PFV climb tree-DO

b. sai Ø ya tsinko 'ya'ya-n itatuwa.
then SUBJ-Ø he-PFV pluck children-of trees-DO
c. Da Ø ya tsinko **wadannan** 'ya-n
when SUBJ-Ø he-PFV pluck **DEM** children-of
itatuwa,
trees-DO
d. sai Ø ya sanya Ø a ciki-n kwando.
then SUBJ-Ø he-PFV put DO-Ø at inside-of basket

(Speaker 2)

'After he had climbed the tree he plucked some fruits. When he had plucked **these fruits** he put (them) in a basket.'

where the demonstrative-anchored inanimate direct object **wadannan 'yan itatuwa** 'these fruits' in (23c) looks back only a single clause to its initial mention, and is, of course, one of the pivotal props around which the whole Pear Story is woven.

Finally in this regard, we note some examples of referents anchored with a determiner shortly after first mention. Excerpt (24) illustrates a human subject from one of the Pear Film stories:

(24) a. Shikenan sai Ø ya yi sa'a,
OK then SUBJ-Ø he-PFV do luck
b. ga wasu yara sun zo daidai
behold IS boys-SUBJ they-PFV come exactly
guri-n,
place-DET-LOC
c. Ø su-na wasa.
SUBJ-Ø they-IMPFV play
d. Sai **yara-n** suka zo
then **boys-DET-SUBJ** they-PFV come
e. Ø suka taimake shi...
SUBJ-Ø they-PFV help him-DO PRO

(Speaker 2)

'OK he was lucky, some kids came right to the place, playing around. Then **the kids** came and helped him...'

where the determiner-anchored human subject NP **yaran** 'the kids' in (24d) looks back 1 clause to its previous (zero) control, and 2 clauses (24b) to its initial mention in the story.

In (25) the suffixal determiner is used to anchor a recently-introduced inanimate referent:

(25) a. Ø Su-na nan zaune,
SUBJ-Ø they-IMPFV there-LOC sit-STAT
b. sai ga cinya-r mutum ta fadɔ
then behold thigh-of man-SUBJ it-PFV fall
kasa tim, har da jini.
to the ground-LOC bump even with blood-ASSOC
c. Yusha'u ya dubi **cinya-r,**
Yusha'u-SUBJ he-PFV look at **thigh-DET-DO**
d. Ø ya ce...
SUBJ-Ø he-PFV say

(Imam 1970:23)

'They were sitting there when a man's thigh fell to the ground with a bump, blood and all. Yusha'u looked at **the thigh** and said...'

The determiner-anchored NP in question -- the direct object **cinyar** 'the thigh' in (25c) -- has a look-back of only 1 clause, again to its first discourse mention in (25b), and persists as the inanimate topic of the immediate episode.

Finally, from another Pear Film narrative we have:

(26) a. ...sai kuma ga wata yarinya ta taho
then and PRESENT IS girl-SUBJ she-PFV come

daga wani wuri.
from IS place

b. Lokacin da Ø suka zo
when SUBJ-Ø they-PFV come

c. Ø za-su gifta,
SUBJ-Ø they-FUT pass by

d. Ø sun gifta juna,
SUBJ-Ø they-PFV pass each other-DO

e. Ø ya-na can
SUBJ-Ø he-IMPFV there-LOC

f. Ø ya-na kallo-n **yarinya-r...**
SUBJ-Ø he-IMPFV watch-of-VN **girl-DET**

(Speaker 1)

'...then a (certain) girl appeared from somewhere. When they (the boy and the girl) came, they were about to pass on, they had passed each other by, and there he (the boy) was, looking at the girl...'

where the determiner-coded NP **yarinyar** 'the girl' in (26f) follows upon first lexical mention in (26a). Again, there is no sense in which the putative selection of a less overt pronoun would have created potential ambiguity in this context.

4.2.2. Coding of culturally-prominent referents

Close investigation of the corpus revealed a tendency for human discourse-participants ranked high within the socio-cultural Hausa hierarchy to have a relatively strong hold on full nominal status, i.e. there were contexts in which an NP was exploited in preference to what would have been an equally unambiguous nonovert form such as a pronoun. For example, one of the stories in Imam (1970) revolves around two (sets of) characters -- 'an

emir' and 'three young men' -- all of them equally central to the unfolding plot. However, whereas 23 out of a total 54 (42.6%) references to 'the emir' (atop the traditional Hausa hierarchy) are achieved with a coreferential NP of some kind, only 3 out of 46 (6.6%) of the references to the three young men were accorded full nominal status. Furthermore, the average look-back for all lexical NP references to 'the emir' was only 6.9 clauses, compared with 15.7 for 'the young men', thus ruling out the possibility of any cognitive explanation relating to the number of intervening clauses. Excerpts (27-28) illustrate the phenomenon in question:

(27) a. Ran nan Ø su-na zaune da
Day DEM SUBJ-Ø they-IMPFV sit-STAT with
sarki,
emir-ASSOC

b. sai wani maroki ya zo
then IS praise singer-SUBJ he-PFV come

c. Ø ya-na ta bunkasa **sarki** da
SUBJ-Ø he-IMPFV keep on flatter **emir-DO** with
kirari,
epithet-ASSOC

d. har **sarki** ya shiga ciki-n
until **emir-SUBJ** he-PFV enter inside-of
abi-n da ya-ke fadi...
thing-DET REL he-IMPFV say-VN

(Imam 1970:9)

'One day they were sitting with the emir when a praise singer came and was flattering **the emir** with epithets, so much so that **the emir** took notice of what he was saying...'

The first nominal reference to 'the emir' -- the associative N sarki in (27a) -- has a look-back of 5 clauses to

its previous control -- also a full nominal -- and the fact that (27a) signals the beginning of a new (paragraph-marked) episode clearly has something to do with the choice encountered here (cf. Chapter 5). Observe, however, that the next two mentions of the same referent in (27c-d) are also encoded with the same full nominal, even though only 2 and 1 clause boundaries respectively intervene between the two successive mentions. It could be argued, of course, that the full subject nominal mention **sarki** 'the emir' in (27d) is here activated by the subject-switching which has taken place, and indeed I shall document numerous such cases in Chapter 5. No such explanation is available, however, for the direct object N **sarki** in (27c), and it is this type of unusually heavy coding, i.e. a full lexical N in preference, for example, to an equally unambiguous pronoun, which is sometimes a feature of such (culturally-determined) high-ranking referents, and which would be unusual outside this restricted domain.

In (28) we have a similar example from one of the spoken narratives:

(28) a. ...Ø suka zo ciki-n gida-n sarki.
...SUBJ-Ø they-PFV come inside-of palace-of emir
b. Ø Suka zo gida-n sarki,
SUBJ-Ø they-PFV come palace-of emir
c. Ø ba su tad da **sarki** ba,
SUBJ-Ø NEG they-PFV find **emir-ASSOC** NEG
d. **sarki** kuwa ya ce...
emir-SUBJ and he-PFV say

(Speaker 3)

'...they entered the emir's palace. They came to the emir's palace, (but) they didn't find **the emir**, and **the emir** said...'

The speaker makes reference to 'the emir' with the full possessor nominal sarki in (28a), repeating the whole clause almost verbatim in (28b). Clauses (28c) and (28d) also contain full nominal mentions of 'the emir', and as was the case in (27c) it is the nonsubject occurrence of another full coreferential nominal **sarki** in (28c) which is again slightly idiosyncratic, with subject-switching probably responsible again for the nominal mention in (28d) as in (27d).[23]

5. Summary

The foregoing analysis has attempted to account for some, at least, of the varying lexical NP forms selected to specify definite referents, and against a background of various conditioning factors. The findings are in many ways incomplete and exploratory, and it is clear that because we have been dealing with a finely-graded referential domain, the numerically-based differences in distribution are certainly not as striking as those noted for the more clearcut binary choices between the Indefinite Specifier and zero-marking of indefinites (Chapter 2), and

the pronoun:zero anaphor opposition for direct object arguments (Chapter 3). The facts thus limit us to the generalization that, all things being equal, there will be circumstances in which the presence, or absence, of a given NP reference-type will be likely, optional, or unlikely. The next, and final chapter considers the impact of two further discourse-based factors -- "subject-switching" and the intervention of episode-boundaries.

Notes to Chapter 4

[1]Despite the admitted deficiencies of this study, it remains the only serious attempt, within the context of Chadic studies, to make sense of the formal distribution of these referential options in their wider discourse context. Schuh's (1977) description of the determiner system in the related Bade/Ngizim language group, for example, is from a restrictively sentence-based perspective, and, as noted in Hopper and Thompson (1983:35), some of the unexplained facts could only be elucidated via scrutiny of a wider corpus.

[2]A more comprehensive report would also include data on such additional reference-types as adjective + N, N + restrictive relative clause etc. However, insufficient tokens rendered such an enterprise at present impossible.

[3]Kirsner (1979), in his contrastive study of the two Modern Dutch demonstrative adjectives "deze" and "die", shows a direct correlation between the high deixis-signaling deze and two dimensions of salience -- humanness and subjecthood. Turning to the written Hausa texts where a reasonably large token count is available, if we restrict our attention exclusively to demonstrative- and determiner-equipped nominals, we encounter the following distributions with respect to these two factors:

	Human		Nonhuman		Subject		Nonsubject	
N + DEM	78/142	55.0%	34/77	44.2%	65/116	56.0%	47/103	45.6%
N + DET	64/142	45.0%	43/77	55.8%	51/116	44.0%	56/103	54.4%

The skewings for the two deictics, whilst not as conclusive as Kirsner's Dutch figures, are generally in the right direction: human referents tend to attract high deixis demonstratives (55.0%) more frequently than determiners (45.0%); and conversely, nonhuman referents favor intermediate deixis-coding determiners (55.8%) over demonstratives (44.2%). And with respect to the variable of subjecthood, subject referents are assigned a proportionately higher percentage of demonstratives (56.0%) than determiners (44.0%); and conversely, there is a preference

for determiner-marking of nonsubjects -- 54.4%, compared with 45.6% for the demonstrative category.

In a later treatment of basically the same problem, Kirsner (p.c.) appeals to the notion of "re-chunking" in order to explain some of the uses of the HIGH DEIXIS demonstrative in Dutch. This process is viewed as correlating with "'retrievals' requiring more interpretation and inference on the hearer's part", and has to do with reintroduction of a referent with a morphologically-related form as opposed to an exact lexical repetition, e.g. rechunking a referent from an 'activity' to a 'thing'. This procedure, whilst not without its merits for the Modern Dutch facts, is not applicable to Hausa data since the vast majority of tokens denoted only basic, concrete referents.

[4]Recall that intervening direct speech quotations were included in the clause counts for the written narratives, and that there were no cases of direct quotations in the spoken texts, a difference which will contribute to the numerical disparities for the two narrative-types in Table 4.2.

[5]Clancy's corresponding data from the Japanese Pear Film accounts show nominal reference to be used over shorter discourse gaps in general.

[6]For the purposes of the immediate analysis, low tones on the demonstrative elements are indicated with a grave (`) accent over the tone-bearing vowel, falling tones with a circumflex (^), and high tones are simply left unmarked. Hausa also possesses a tonally distinct nan-marked demonstrative set, where the tonal alternations in the postnominal array are related to the tone on the final syllable of the preceding nominal. The occurring forms are:

Prenominal: wannàn (masc./fem.sg.)/wadànnân (pl.)

Postnominal: -n nàn/nân (masc.sg./pl.)/-r nàn/nân (fem.sg.)

These particular forms are used to map referents which are either physically visible or abstract and non-visible to the encoder or decoder, e.g. **wannàn** yaro 'this boy' [+visible], vs. **wannàn** gaskiya ne 'this/that is true' [-visible], where the [-visible] anaphoric **wannàn** is typically used to confirm or summarize a prior proposition. Although these demonstratives are especially common in direct speech contexts involving bilateral exchanges, examples from indirect reportative narrative are available, e.g. ...ya kai **wannàn** gari ya wuce, ya kai **wannàn** ya

wuce... (Imam 1970:58) '...he would reach this town and pass by, and would reach this/that town and pass by...'. It may be that the use of the high-low tone **wannàn** demonstrative, in contrast to the low-high tone **wànnan** form normally used as the referential operator in narrative discourse, is being manipulated to signal an encoder-oriented strategy. That is, the encoder is seen as actually moving into the world of the narrative itself, and describes events as if s/he were participating in them (cf. Chafe 1980a, Clancy 1980:146ff., and Tannen 1980 for discussions of the impact of such "world shifts" on referential choice). Finally, Russell Schuh (p.c.) has also pointed out that there is a certain amount of speaker/dialectal variation with regard to use of all the above demonstratives. I checked a sample of the demonstratives in the written corpus with two (Standard) Kano Hausa speakers, both of whom regularly supplied the tonally distinct forms discussed in this section.

[7]This pronominal function has a much lower text-frequency than the lexical NP-qualifying adjectival usage considered here, and I am restricting this preliminary analysis to the latter usage (cf. Kirsner 1979:356).

[8]The full form can in fact be postposed, but no examples were found in the corpus. For the purposes of the present investigation, I am treating both the preposed and postposed variants as signaling the same meanings basically. I am aware, however, of the kind of work that Klein-Andreu (1983b) has done on Spanish adjectives, where she demonstrates that there is an observable correlation between position and differential meaning.

[9]Hausa also possesses a can-demonstrative set characterized by exactly the same tonal alternations as the nan-set, though an insufficient token count has excluded them from present consideration. The high-tone can-marked variants differ from their nan counterparts in that they are used to specify only non-visible, purely referential items, and my impression is that they serve, in general, to denote referents in contexts which are highly discontinuous, a characterization which would be consonant with the claim in Jaggar (1983:423) that their basic function is one of "mapping objects which are more remote-distal in a temporal or spatial sense." Along with some of the empirical and interpretative issues mentioned above, this represents another potentially fruitful line of research within the domain of demonstrative reference in Hausa.

[10]Russell Schuh (p.c.) has reminded me that Abraham (1959:53-55) discusses some facets of the context-dependent visible:referential distinction encoded by these items, but he makes no mention of this encoder-proximate function.

[11]The NP is also modified with an independent pronominal possessor nasa 'of his'. However, the presence of a possessor argument with such kin-terms is virtually obligatory in Hausa (cf. Jaggar 1981).

[12]When working with discourse intervals of the size exemplified in this section, I am aware that the intervention of one or more episode-boundaries is almost inevitable, and as will be shown in Chapter 5, such breaks in discourse-continuity do have an observable impact upon referential choice. My main concern at this stage, however, is to illustrate the kinds of discontinuous contexts which typically precipitate use of the various categories, thereby introducing these same categories to the reader.

[13]The internal evidence for this proposed derivation relates to the fact that selection of either a determiner or postposed high-tone demonstrative triggers identical tonal changes on the final syllable of the preceding noun, though Russell Schuh (p.c.) has suggested that the determiner was originally the demonstrative itself, and the postposed nan/can variants of Modern Hausa represent an accretion of some kind. From a comparative viewpoint, as noted in Givón (1982:25), "the development of the remote demonstrative as the definite article is probably the most widespread process, cross-linguistically, of obtaining definite articles" -- cf. too Greenberg (1977, 1978). In Jaggar (1983) -- following the terminology generally used in Chadic language studies (cf. Schuh 1977:26ff., 1983) -- I used the term "Previous Reference Marker" to describe this referential suffix. I now prefer to use the label "determiner", mainly in order to make a distinction between this morpheme and the demonstrative, both of which, strictly speaking, encode entities which are "previously referred to".

[14]This discourse stretch was the largest I encountered for any NP configuration, and it is perhaps a little unusual in that one might have expected a high deixis-signaling demonstrative in such a strongly disruptive context. I think that the semantic class of the NP here -- a specific relational kin-term -- probably accounts for the choice of the less dense determiner-coding. Chapter 5 looks at some of the coding conventions associated with kin-terms.

[15]Interestingly, even comparison of Imam (1970) with earlier editions of the same volume reveals a similar trend.

[16]As Russell Schuh (p.c.) has correctly reminded me, however, proper name + determiner combinations are perfectly acceptable in Hausa, e.g. Audu-n (Audu-DET) 'The Audu (we all know)'.

[17]Unlike some of the foregoing examples, where the issue of potential ambiguity is perhaps less clearcut, this appears to be an obvious case of an explicit device activated by potential confusion of referents.

[18]The sense in which I am using the term "prop" differs slightly from Du Bois' (1980b:269ff.) characterization in that my definition allows for such "props" to be deployable and so exhibit continuing identity throughout the discourse.

[19]See Wald (1983:100-101) for a discussion of some interesting parallels from Chaucerian (Middle) English, where unstressed anaphoric 'this' is used to code second lexical mention of a newly-introduced referent, especially when the referent in question occupies the subject position.

[20]Notice my English gloss 'one day' for the demonstrative-marked ran nan in (19a), literally 'that day'. I think the reason for the collocational use of the demonstrative here may relate to the fact that this is another encoder-centered strategy -- the writer is aware of the particular 'day' itself, and is here depicting the temporal switch from a perspective internal to the narrative and consonant with his own knowledge.

[21]Clancy (1980:154ff.), in the context of a similar discussion of the Japanese Pear Film narratives, talks in terms of a "subtle transition from "introduction" to "action" in the story line", and provides English glosses which suggest the presence of demonstratives being manipulated to achieve the same communicative goals in the Japanese. Interestingly, she remarks that no comparable examples were encountered in the English accounts. And Robert Kirsner (p.c.) has drawn my attention to Hinds' (1977) discussion of the use, in English obituaries, of full NP's in preference to pronouns.

[22]The fact that Hausa has grammatical gender -- third (and second) person singular masculine/feminine pronouns are formally distinguished -- is clearly relevant to the

issue of potential ambiguity.

[23]Paul Schachter (p.c.) has also suggested that in English, where no tendency to associate lexical noun references with rank is apparent, there might be a similar preference for lexical mention following possessor reference, i.e. 'They came to the emir's palace but didn't find **the emir**', rather than 'They came to the emir's palace but didn't find **him**'.

CHAPTER 5

REFERENTIAL CHOICE, SUBJECT-SWITCHING/-PRESERVATION AND LARGER DISCOURSE BREAKS

1. Introduction

In the preceding chapter we discussed a number of discourse-determined factors which typically persuade an encoder to revert to full nominal coreference for an entity already introduced into the discourse, i.e. use of a full NP -- simple noun, noun with determiner or demonstrative -- to refer to a pre-established, identifiable referent. In this final chapter, we take a more detailed look at the interrelationship between ongoing referential choice and three structural phenomena. In section 2, I present an introductory overview of the problems to be addressed. Sections 3 and 4 then consider, respectively, the influence both subject-preservation and subject-switching exert upon the encoder's selection of nonovert elliptical (3.1, 4.1) versus overt nominal (3.2, 4.2) reference, and we shall investigate the relations between the above factors with respect to the variable of intervening sentential boundaries. Section 5 then focuses upon the effects of some higher-level discourse breaks on

referential choice, including the occurrence of paragraph-marked boundaries and fragments of quoted direct speech. The kinds of questions we shall attempt to answer are: 'Under what general conditions might we expect to encounter a particular referential form? Are we able to suggest an analysis which can shed light on the generalized norms of distribution?'

Some of the issues I address have already been touched upon in previous chapters, and the basic aim of this closing chapter is to pick up some of the loose ends by providing a more comprehensive and rigorous treatment of these issues. As noted in Chapter 3, moreover, it should be borne in mind that, with regard to the domain of NP coreference, we are handling a finely-graded continuum which plots a relatively wide range of non-discrete coding-points, e.g. simple noun, noun + determiner, noun + demonstrative, many of whose functions often seem to overlap in the context of live discourse. For this reason, the quantified regularities expressed in this and the preceding chapter are necessarily more limited than those stated in Chapters 2 and 3, for the analyses proposed in these earlier chapters the descriptive advantage of dealing with reasonably clearcut, binary choices -- Indefinite Specifier vs. zero, and direct object pronoun vs. zero anaphora -- which allowed more conclusive statistical skewings.

Such reservations notwithstanding, the following investigation proceeds, as ever, in the spirit which dictates that we should "make use of statistical generalizations across a corpus of texts, in order to define the otherwise elusive function of certain forms at the level of discourse" (Du Bois 1981:258).

2. Subject-preservation and subject-switching: a quantitative overview

The data for the analysis to follow derive from a count of the subject arguments of <u>all</u> the clauses, main and subordinate, in the first seven stories of Imam (1970:6-42) -- over 1,300 tokens in all -- and the four oral narratives available -- a little over 200 type-tokens.[1]

Recall, first of all, that in the preverbal subject position, the only two referential strategies used with any significant cross-text frequency are the "extremes" of full NP's and ellipsis, with subject ellipsis constituting the canonical form of nonovert anaphoric reference.[2] For the reader's convenience, these two reference-types are again illustrated in (2) and (3):

(2) **Yarinya** ta fita — Lexical noun subject
girl-SUBJ she-PFV go out
'**The girl** went out'

(3) **∅** Ta fita — Subject-ellipsis
SUBJ-∅ she-PFV go out
'∅ (She/the girl) went out'

In examples (2-3), obligatory subject-agreement morphology is incorporated in the preverbal auxiliary ta (third person feminine singular perfective). The agreement system in Hausa is thus somewhat analogous to the situation in Spanish (cf. Bentivoglio 1983), and covers a referential domain which is mapped by "true" zero-anaphora of the subject in languages such as Japanese (Clancy 1980; Hinds 1983), i.e. where there is no formal marking on the verb (or auxiliary) to provide any overt clues to the identity of the ellipted subject.

Figures 5.1 and 5.2 (cf. Appendix IX) now present written and oral narrative data relating to the distributional percentages of elliptical vs. nominal reference as used for preserving and switching subjects, both within and across sentence boundaries.[3]

FIGURE 5.1. Written Narrative Percentages of NP's and Ellipsis Selected for Preserving and Switching Subject Reference within and across Sentences.

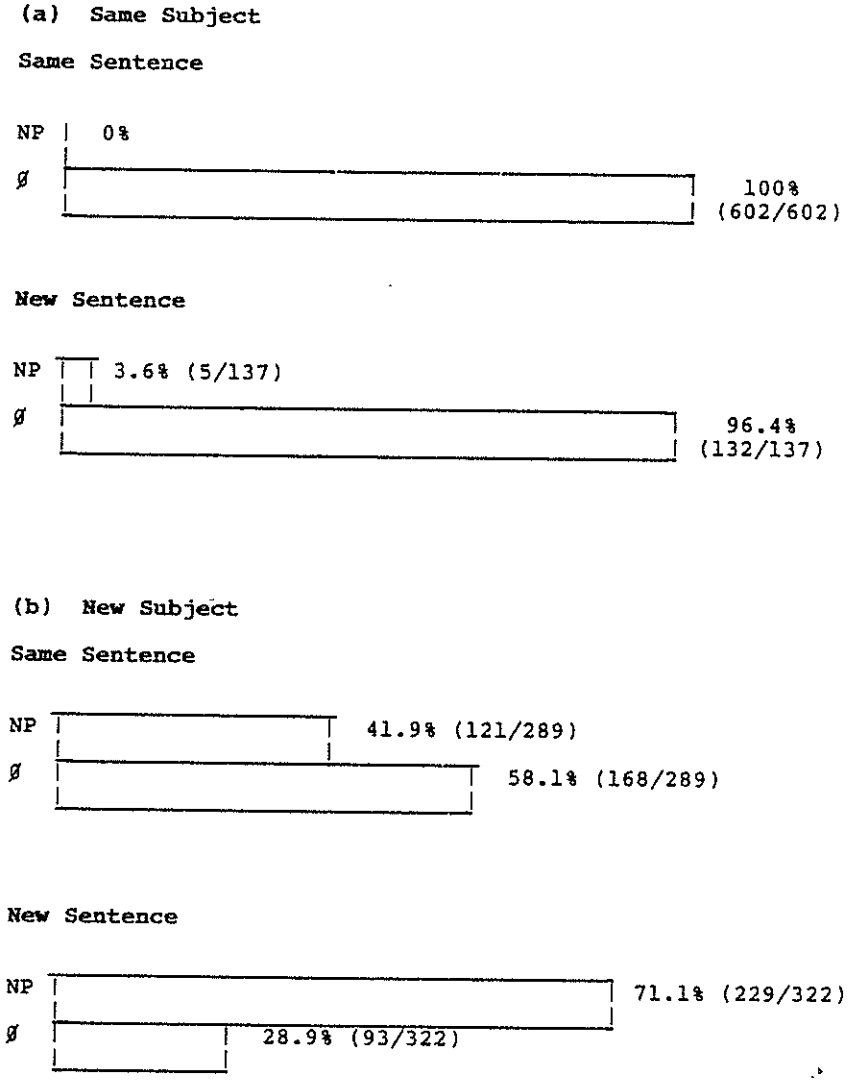
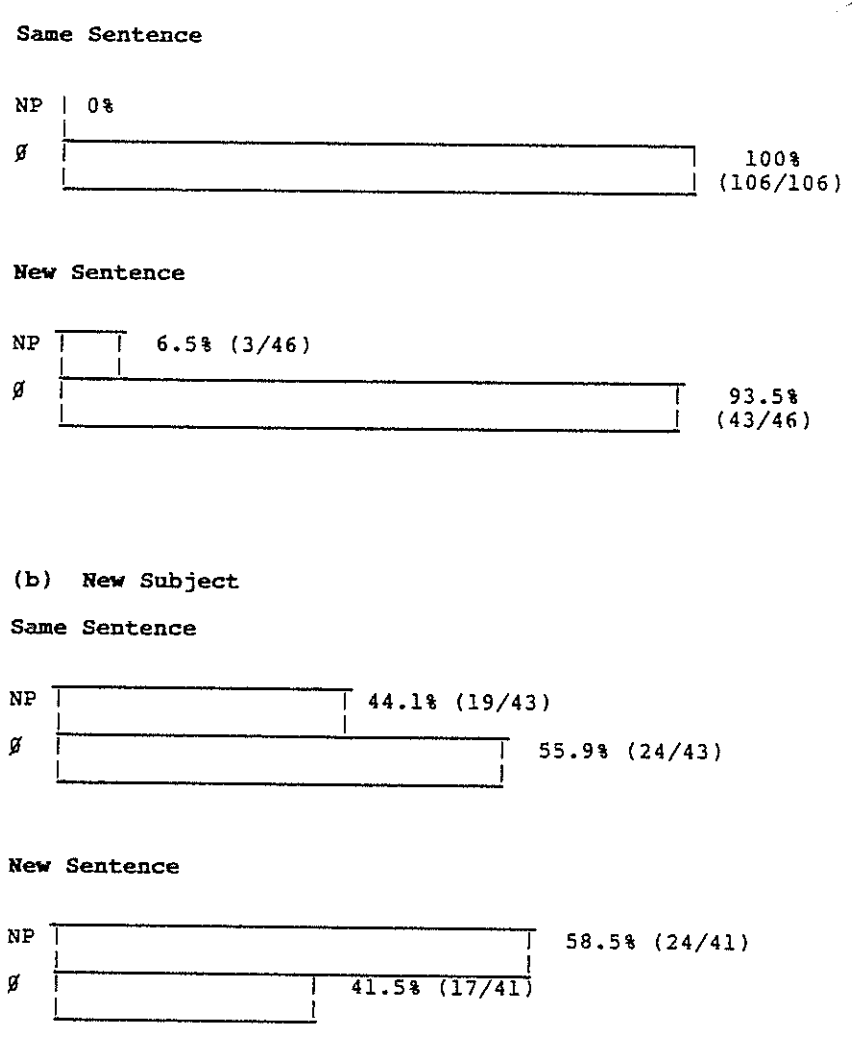


FIGURE 5.2. Oral Narrative Percentages of NP's and Ellipsis Selected for Preserving and Switching Subject Reference within and across Sentences.



It is clear from the percentages given in Figures 5.1 and 5.2 that a discernible interpredictability exists between referential choice, preservation vs. changing of the subject argument, and the partitioning of discourse chunks into sentences. Sections 3 and 4 now provide an elaboration of these observations.

3. Subject-preservation and referential choice

Inspection of the data in Figures 5.1 and 5.2 shows that ellipsis was clearly the preferred option, regardless of whether a sentence-boundary had in fact intervened.

3.1. Subject-ellipsis

In Chapter 3, it was demonstrated that zero-anaphora in the subject position occurred most typically in tightly-integrated and highly continuous discourse contexts -- measured, that is, in terms of the number of intervening clauses and maximally similar potential confusers. The data in the present section on subject-maintenance, and in section 4 on subject-switching, are intended to show that even though subject-ellipsis can in fact be used in more disruptive environments, there are still conditions which act to constrain its exploitation.

Looking at Figures 5.1(a) and 5.2(a), we note that

when the subject was maintained within a sentence-boundary, subject arguments were, without exception, coded with zero-anaphora -- values of 100% in both the written and spoken texts. Furthermore, even if a sentence break did in fact intervene, subject-ellipsis remained the overwhelming norm -- 96.4% (written) and a remarkably similar 93.5% (spoken). These values may be compared with those provided by Clancy (1980:162ff., 202) for ellipted same-subjects -- 37.5% (same sentence) and 8.0% (new sentence) in English, and 88.0% (same sentence) vs. 83.0% (new sentence) in Japanese. Japanese and Hausa are structurally analogous in that, unlike English, they present speakers with essentially only one attenuate referential device in the subject position, i.e. ellipsis, though the language-specific definitions of this function are, as we have noted, not identical. In same-subject environments, where constant nominal coreference, i.e. use of a coreferential NP to denote a definite referent, would obviously be unnecessary and undesirable, even at the beginning of new sentences, ellipsis remains the only alternative to be exploited.

Although there are clear distributional similarities for ellipsis in Hausa and Japanese, it is worth noting that the respective percentages are not identical -- the strategy of subject-$\emptyset$ in Hausa is less constrained in its distribution, both intra- and inter-sententially, than its

Japanese counterpart, a difference which no doubt derives from the fact that agreement on the Hausa auxiliary element conveys crucial gender-number information regarding the identity of the ellipted subject. Fragment (4) illustrates how the subject-$\emptyset$ device may operate both within and across sentence-boundaries with relative freedom:

(4) a. $\emptyset$ Suka shirya,
SUBJ-$\emptyset$ they-PFV get ready
b. $\emptyset$ suka dunguma zuwa Bila.
SUBJ-$\emptyset$ they-PFV troop off to Bila-LOC
c. Da $\emptyset$ suka isa
when SUBJ-$\emptyset$ they-PFV arrive
d. $\emptyset$ suka nemi
SUBJ-$\emptyset$ they-PFV seek
e. $\emptyset$ su ga sarki,
SUBJ-$\emptyset$ they-SUBJCTV see emir-DO
f. aka yi musu iso.
IMPERS-PFV do to-them-IO PRO arrival
g. $\emptyset$ Suka tafi
SUBJ-$\emptyset$ they-PFV go
h. $\emptyset$ suka fadi
SUBJ-$\emptyset$ they-PFV fall down
i. $\emptyset$ suka yi gaisuwa.
SUBJ-$\emptyset$ they-PFV do greeting

(Imam 1970:7)

'$\emptyset$ (They/the three brothers) got ready and $\emptyset$ trooped off to Bila.
When $\emptyset$ (they) arrived $\emptyset$ (they) sought $\emptyset$ to see the emir and their arrival was announced.
$\emptyset$ (They) went in, $\emptyset$ prostrated themselves and $\emptyset$ greeted (the emir).'

In (4) the (perfective third person plural) subject-$\emptyset$ strategy <u>suka</u> is used both within sentences -- (4b,d,e,h,i)[4] -- across sentence boundaries -- (4a,c,g) -- and also across an orthographic paragraph-boundary in (4c) -- cf. section 5 for details.

Excerpt (5) is taken from the beginning of a Pear Film account:

(5) a. ...Ø ya je
SUBJ-Ø he-PFV go
b. Ø ya-na tsinka-r mangwaro-n-shi.
SUBJ-Ø he-IMPFV pick-of mangoes-of-his
c. Ø Ya-na da kwanduna guća uku.
SUBJ-Ø he-IMPFV with baskets unit three-ASSOC
d. Shikenan, Ø ya cika kwando
OK SUBJ-Ø he-PFV fill basket
guda biyu,
unit two-DO
e. Ø ya hau ka-n mangwaro-n,
SUBJ-Ø he-PFV climb top-of mango tree-DET
f. Ø zai je
SUBJ-Ø he-FUT go
g. don Ø ya tsinko mangwaro-n da
so that SUBJ-Ø he-SUBJCTV pluck mango-DET REL
Ø zai cika kwando-n-shi na ƙarshe...
SUBJ-Ø he-FUT fill basket-of-his of last-DO

(Speaker 1)

'...Ø (he) had gone and Ø was picking his mangoes.
Ø (He) had three baskets.
That was that, Ø (he) had filled two baskets and Ø had climbed the mango tree and Ø was about to go and Ø pick mangoes so that Ø (he) could fill his last basket...'

In (5), zero-anaphoric reference to the subject -- 'the mango-man' -- is preserved within and across several sentence-boundaries, and in (5d), across a paragraph-marked episode-boundary (cf. section 5).

3.2. Nominal coreference

When we turn to consideration of the most explicit device available to speakers, Figures 5.1(a) and 5.2(a)

show that noun phrases were especially favored in cases of subject-switching and when a sentence-boundary had been traversed. The same data also reveal that no NP-coded same subject/same sentence tokens were encountered anywhere in the corpora investigated, a finding which certainly accords with intuitions on this matter. It seems intuitively reasonable to assume too that even the occurrence of a coreferential subject nominal in a subsequent new sentence would be highly unusual, and indeed this turns out to be the case in reality -- only 5 out of 137 (3.6%) attested in the written, and 3 out of 46 (6.5%) recorded in the spoken stories. Most of these cases, it should be noted, also coincided with major episode-boundaries, detailed treatment of which is presented in section 5. Several did not co-occur at such break-points, however, and I would like to conclude this section on subject-preservation with exemplification and brief discussion of one of these cases. Consider, therefore, fragment (6):

(6) a. Isa ya tashi
Isa-SUBJ he-PFV get up
b. Ø ya ɗauki ƙota-n nan
SUBJ-Ø he-PFV take handle-DET DEM-DO
c. Ø ya yi ta bugu-n Buzu
SUBJ-Ø he continue beating-of VN Tuareg
d. har Ø ya faɗi ƙasa.
until SUBJ-Ø he-PFV fall to ground-LOC
e. Buzu-n dai bai ce i ba,
Tuareg-DET-SUBJ however NEG-he-PFV say yes NEG
balle a'a.
let alone no

(Imam 1970:12)

'Isa got up, picked up that handle and kept on beating the Tuareg until he fell to the ground. **The Tuareg**, however, didn't say a word.'

Although the new sentence (6e) does not appear to constitute the kind of radical shift we shall be looking at later, observe the use of the determiner-supported same-subject NP Buzun 'the Tuareg'. I think several factors may be responsible for this referential choice. Whereas the proper name subject Isa is the highly agential controller of the voluntary actions verbalized in clauses (6a-c), this is not the case with the new subject in (6d), i.e. 'the Tuareg'. It is as if, in this case at least, inadvertent control over an involuntary, non-transitive action like "falling" is not sufficient to establish the switch-subject in the discourse "driving seat"; further explicit specification is deemed necessary, therefore, and is duly supplied in the form of a determiner-equipped NP.[5] Paul Schachter (p.c.) has also noted that while Isa is a main-clause subject, Buzu is only a subordinate-clause subject and that since they are maximally similar arguments, putative use of an ellipted subject in (6e) would create referent-ambiguity. Finally, it may be of relevance to note that the topicalized subject NP Buzun in (6e) is followed by the modal particle dai, translated in this context as 'on the other hand, however' by the two

Hausa-speakers I consulted; perhaps, therefore, use of this counterexpectational particle creates a sufficiently disruptive environment to induce overt NP coreference.[6]

4. Subject-switching and referential choice

Our attention now turns to an especially prominent type of interference which regularly produces cases of potential ambiguity of reference and so motivates use of explicit coreference. This type of ambiguity is encountered in environments where a referent is reintroduced in the subject slot, following a clause or sentence containing a different subject argument, i.e. in cases of **"subject-switching"**.[7]

Comparison of Figures 5.1(a) and 5.2(a) with Figures 5.1(b) and 5.2(b) shows quite clearly that subject-switching elicits quite different referential choices, resulting in greatly increased percentage-occurrences for NP's and lower percentages for subject-ellipsis. The data also demonstrate that these ongoing choices are sensitive to the presence, or not, of sentence boundaries.

4.1. Subject-ellipsis

Looking, firstly, at the more substantial body of written data in Figure 5.1(b), we see that ellipsis re-

mains the favored device even at points of subject-switching, with the proviso, however, that the switch occurs within the same sentence -- 58.1% (same sentence), compared with 28.9% across sentence boundaries. And with respect to nominal coreference, we have the converse distribution -- 41.9% coding switch-subjects within the same sentence, contrasting with 71.1% across sentences. In this regard, the respective distribution of the subject-ellipsis in contrast to subject NP's is closer to the subject pronoun:subject NP distribution reported by Clancy (1980) for English -- 53.1% (same sentence) subject pronouns vs. 34.1% (new sentence), and 65.4% new sentence NP's vs. 42.1% in the same sentence. In the Japanese Pear Film data, on the other hand, the numerical disparities between subject-ellipsis and NP coreference in switch-subject environments are more extensive, due to the fact that, unlike Hausa, subject-ellipsis in Japanese preserves no information about the controlling subject argument, and so the potential for referent ambiguity is necessarily increased. With respect to Hausa, the more noteworthy examples of elliptical subject-switching involve contexts in which maximally similar arguments competed, at least hypothetically, for coreferential control of aux-agreement (most frequently third person singular masculine in the corpus examined), and I shall restrict my examples and remarks to this category.[8] Excerpt (7)

illustrates ellipsis-coded subject-switching within the same sentence:

(7) a. Kowanne-n-su ø ya yi musu addu'a,
Each-of-them SUBJ-ø he-PFV do to them-IO PRO prayer
b. ø suka ce
SUBJ-ø they-PFV say
c. yadda ø ya ji ƙa-n-su haka,
just as SUBJ-ø he-PFV feel mercy-of-them thus
d. Allah ya ji ƙa-n-sa
Allah-SUBJ he-SUBJCTV feel mercy-of-him
shi kuma,
he-IND PRO too
e. ø ya saka masa fiye da
SUBJ-ø he-SUBJCTV give to him-IO more than
abi-n da ø ya ba su.
thing-DET REL SUBJ-ø he-PFV give them

(Imam 1970:10-11)

'And to each of them ø (he/Isa) gave a prayer, and ø (they) said that just as ø (he) had shown mercy on them in this way, Allah would show mercy on him too, and ø (he/Allah) would give him more than ø (he/Isa) had given them.'

In the context immediately preceding (7), the proper name referent Isa has been distributing largesse to all the unfortunates in the area, and this knowledge serves to provide clarification of the subject referent of the relative clause ellipsis device in (7e).

Fragment (8) illustrates ellipsis-coded subject-switching both within and across sentence-boundaries:

(8) a. Baƙo tsammani ya-ke
stranger-SUBJ thinking he-IMPFV
b. Kalala ø-na nufi-n
Kalala-SUBJ CONC-ø-IMPFV mean-of
c. kunne-n-sa daya kaɗai ø ya-ke so
ear-of-his one only SUBJ-ø he-IMPFV want-VN

d. Ø ya yanka.
SUBJ-Ø he-SUBJCTV cut off
e. Ø Ya waiwaya
SUBJ-Ø he-PFV turn round
f. sai Ø ya ga
then SUBJ-Ø he-PFV see
g. Ø ya taso masa da wuka
SUBJ-Ø he-PFV come to him-IO PRO with knife
tsirara a hannu.
naked-ASSOC in hand-LOC

(Imam 1970:19)

'The stranger thought Kalala meant that Ø (he/Kalala) wanted to Ø cut off one of his ears.
Ø (He/the stranger) turned round and Ø saw that Ø (he/Kalala) had made for him naked knife in hand.'

Again, it is the prior discourse context which provides a natural basis for interpreting the referents of the various elliptical subject-switches in (8c-g) -- we know that it is Kalala who has has picked up a knife and set off in pursuit of 'the stranger'. Consider also excerpt (9):

(9) a. Da Isa ya ga
when Isa-SUBJ he-PFV see
b. Ø ya ciwo ka-n-sa,
SUBJ-Ø he-PFV win over head-of-his-DO
c. Ø ya raka shi gida
SUBJ-Ø he-PFV escort him-DO PRO home-LOC
d. Ø ya dawo.
SUBJ-Ø he-PFV return
e. Ko yaushe Ø ya zo
Whenever SUBJ-Ø he-PFV come
f. Ø ya yi masa aski
SUBJ-Ø he-SUBJCTV do to him-IO PRO shaving-DO
g. sai Ø ya tuna masa da
then SUBJ-Ø he-SUBJCTV remind to him-IO PRO with
wauta-r da Ø ya yi...
stupidity-DET REL SUBJ-Ø he-PFV do

(Imam 1970:14)

'When Isa saw that Ø (he/Isa) had won him (the barber) over, Ø he escorted him home and Ø came

back. And whenever Ø (he/the barber) came Ø to give him (Isa) a shave, Ø (he/Isa) would remind him of the stupid thing Ø (he/the barber) had done...'

There are several reasons why the subject-Ø strategy in the new sentence (9e) could only be interpreted as denoting 'the barber': firstly, the motion verb dawo 'return' in (9d) establishes the locative setting for the actions of the subsequent clauses, i.e. 'Isa's' house; secondly, the lexicosemantics of the verb 'to shave'; and finally, barbers in Hausaland usually visit their customers.

Instances of "genuine" ambiguity, i.e. points at which the decoder -- myself in this case -- experienced some difficulty in immediately interpreting an ellipted referential expression -- were non-existent in reality.[9] Thus, in those cases which did counter the general tendency to interpret a string of elliptical subjects as specifying the same referent, other factors -- most typically prior narrative context and lexicosemantic subcategorization of verbs, but also culturally-determined expectations -- acted to clarify referent identity.

With respect to the oral narrative data summarized in Figure 5.2(b), one feature perhaps worth noting is that the switch-subject new sentence ellipsis:NP percentages are a little closer than was the case in the written narratives -- 58.5% (NP) and 41.5% (ellipsis), compared with 71.1% and 28.9% respectively in the written stories.

This disparity is probably attributable to the fact that in unplanned, informal oral narratives, as opposed to tightly-edited and formal written narratives (cf. Chafe 1979a; Tannen 1981, 1982), the encoder has less opportunity to pre-plan and monitor the exploitation of these referential devices.[10] Despite the absence of conscious editing, however, no processing difficulties emerged in the spoken narratives.

4.2. Nominal coreference

Inspection of the written narrative data in Figure 5.1(b) reveals a stronger preference for coreferential NP's across sentence boundaries (71.1%) than within the same sentence as the preceding different subject (41.9%), with corresponding percentages of 58.5% and 44.1% in the oral narratives (cf. Figure 5.2(b)). The exploitation of lexical coreference at points of subject-switching, whether within or over sentence boundaries, thus represents the strongest and most common motivation for using overt devices over relatively short discourse intervals. Fragments (10-12) nicely illustrate the phenomenon:

(10) a. **Buzu** ya zare wuƙa
Tuareg-SUBJ he-PFV unsheath knife-DO
b. Ø ya daka wa sarki-n ɓarayi,
SUBJ-Ø he-PFV stab to king-of thieves-IO
c. wuƙa ta karye uku.
knife-SUBJ she-PFV break three

d. Sarki-n ɓarayi ya zaburam masa,
King-of thieves-SUBJ he-PFV spring to him-IO PRO
e. **Buzu** ya zare takobi
Tuareg-SUBJ he-PFV unsheath sword-DO
f. Ø ya cira Ø
SUBJ-Ø he-PFV raise DO-Ø
g. Ø zai ɗauke ka-n-sa...
SUBJ-Ø he-FUT take off head-of-his-DO

(Imam 1970:38)

'**The Tuareg** unsheathed a knife, stabbed at the King of the Thieves and the knife broke in three pieces. **The King of the Thieves** leapt at him and **the Tuareg** unsheathed a sword, raised (it) and was about to take his head off...'

The three NP subject references in (10a, d, e) are all crucial to the decoder's comprehension, for use of an inexplicit subject-Ø device at these points would result in true ambiguity -- there are the two maximally similar, third person masculine singular human arguments involved in the interchange of actions -- 'the Tuareg' and 'the King of the Thieves' -- and both could control the actions verbalized in their respective clauses. Fragment (11) is similar:

(11) a. **Abdu** ya lura da iri-n siffa-r-ta
Abdu-SUBJ he-PFV observe kind-of shape-of-her
da kyau, tun daga ƙafafu-n-ta har ya zuwa
well-ADV right from feet-of-her up to
bisa kai.
on head-ADV
b. **Nahi'u Hankali** kuwa bai ko
Nahi'u-the-Careful-SUBJ and NEG-he-PFV even
dube ta ba.
look at her-DO PRO NEG
c. Ashe kuma sa'ad da **Abdu** Ø-ke
Well and when **Abdu-SUBJ** CONC-Ø-IMPFV
lura da iri-n-ta,
observe-VN sort-of-her

d. Halilu Wayo ya kaikaici
Halilu-the-Crafty-SUBJ he-PFV distract
idanu-n-su...
eyes-of-them-DO

(Imam 1970:8)

'Abdu observed her (the bird's) features carefully,
from her feet to the top of her head. And
Nahi'u-the-Careful didn't even look at her. And
when Abdu was observing her features,
Halilu-the-Crafty distracted their attention and...'

where a succession of four proper names, denoting the three 'brothers' in the story, is used to preclude any possible ambiguity.

As a final illustration -- the examples could, of course, be multiplied, but I do not wish to belabor the point -- consider (12):

(12) a. ø Su-na nan zaune,
SUBJ-ø they-IMPFV there sit-STATIVE
b. bakauye ø-na ta tunani-n
villager-SUBJ CONC-ø-IMPFV continue thinking-of
karya-r da ø zai shara
lie-DET REL SUBJ-ø he-FUT tell
c. in an tambaye shi,
if IMPERS-PFV question him-DO PRO
d. sai ga abinci attajiri-n nan
then PRESENT food merchant-DET DEM-SUBJ
ya sa
he-PFV cause
e. wani bara-n-sa ya fara kawo-wa ø.
IS servant-of-his-SUBJ he-PFV begin bring-VN DO-ø
f. Ko da ø ya zo
As soon as SUBJ-ø he-PFV come
g. ø ya ajiye ø,
SUBJ-ø he-PFV put down DO-ø
h. ø ya fita,
SUBJ-ø he-PFV go out
i. sai bakauye-n nan ya ce wa
then villager-DET DEM-SUBJ he-PFV say to
kane-n-sa...
younger brother-of-his-IO

(Imam 1970:32)

'They (the villager and his younger brother) were
sitting there, and the villager was contemplating
the lie he would tell if he was questioned, when
some food appeared -- that merchant had got
one of his servants to start bringing (it). As
soon as he came and put (it) down and went out,
this villager said to his younger brother...'

In (12), the role of grammatical subject is bandied around between several maximally similar arguments, and explicit mention at points of subject-switching -- clauses (12b, d, e, i) -- clearly helps to avoid the potential confusion, momentary or otherwise, which use of ellipsis would otherwise cause.

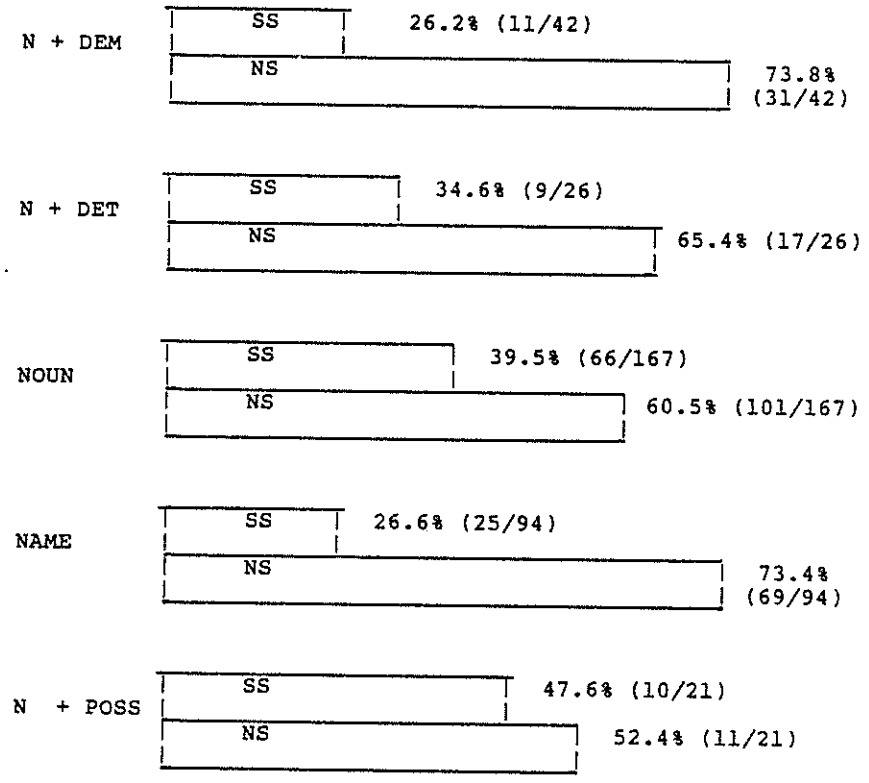
The observant reader will have noticed the following fact about extract (12): the differential coding assigned to the various subject-switch definite NP's exemplified -- simple N(oun) bakauye 'the villager' in (12b); N + demonstrative attajirin nan 'that merchant' in (12d); and N + demonstrative bakauyen nan 'this villager' in (12i) -- an important matter to which we now turn.

Recall the discriminant approach to NP coreference developed in Chapter 4, which attempted to demonstrate an interpredictability between deictic force and nominal categories, especially the configurations N + demonstrative, N + determiner, and simple N. Given the general patterns which emerged from that analysis, an interesting issue worth pursuing is whether there are any analogous

correlations holding between the categories of NP selected to map the discontinuous function of subject-switching across sentence boundaries. Figure 5.3 (cf. Appendix X) summarizes some data from the larger corpus of written narratives.

FIGURE 5.3. Written Narrative Percentages of Differential NP Configurations Chosen for Switching Subjects within and across Sentence Boundaries.

Key: SS = Same Sentence
NS = New Sentence



It seems intuitively plausible to assume that subject-switching at sentential boundaries represents a more radical break in discourse continuity than across sentence-internal clause boundaries, imposing a greater processing burden on the decoder (cf. Chafe 1979b). Given this assumption, moreover, the deixis:coding hypothesis would predict the following relationship: the stronger the deictic force associated with a given NP configuration, the greater will be the tendency to utilize that configuration as a coreferential expression in contexts which involve the crossing of sentence boundaries. The distributional data on the first three NP categories in Figure 5.3, although quite clearly less than conclusive, would seem to suggest a general tendency of this kind. Thus, demonstrative-bearing switch-subject NP's, in keeping with the strong deixis-signaling properties of this category, display the highest cross-sentence percentage -- 73.8% -- followed by the intermediate deixis N + determiner category -- 65.4% -- and finally the weak deixis-mapping class of simple nouns -- 60.5%.

Although the cross-text data summarized in Figure 5.3 seem to point to some possible correlations, to provide category-by-category instantiations of the suggested regularities would, I believe, be misleading. This is due to the fact that, with the possible exception of examples like (10-12), it proved extremely difficult to isolate

nominal choices whose occurrence could be plausibly analyzed as deriving exclusively from a single motivating factor such as avoidance of referent ambiguity at subject-switch points, an analytical dilemma which is, of course, a regular feature of discourse investigations (cf. Clancy 1980:170-71). Furthermore, close scrutiny of all the 350 NP tokens given in Figure 5.3 revealed that a substantial proportion were used following the intervention of one or more episode-boundaries, regardless of whether a maximally similar and so potentially ambiguous subject argument had in fact intervened. Because of this, I believe that a more sensible and illuminating approach would be to consider the distribution of these nominal categories in the context of a wider investigation of the impact which larger discourse breaks have upon referential choice. This is the concern now of the final section.

5. Paragraph-signaled discourse breaks and referential choice

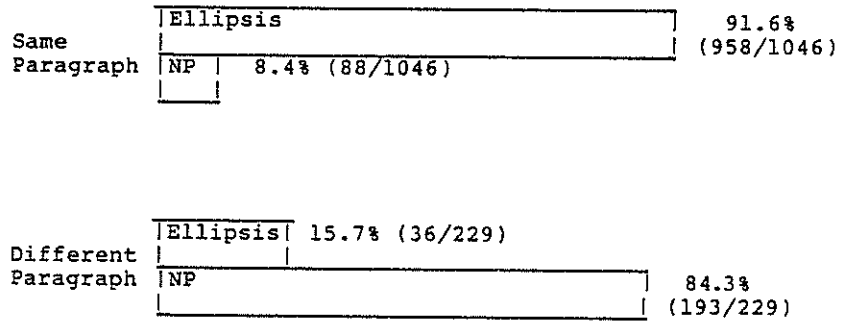
The analysis now formulated is based upon the premise that in written texts, the formal delineation of paragraph boundaries is an important means of partitioning internally-structured stretches of discourse, and that decisions to mark orthographic paragraphs are often related to such discourse-determined factors as the realignment of parti-

cipants, shifts in event-sequences, temporal and locational orientation etc. Indeed, many writers have taken this assumption as given for purposes of text analysis (cf. Grimes 1975:102ff; Hinds 1977, 1978a, 1979:136ff.; Chafe 1979b; Fries 1980; Givón 1981).[11] Such divisions, I am presuming, are simply realistic reflections of the writer's specific goals, visual clues to what s/he considers to be a significant "break-point" in the text, and often coinciding with episodic boundaries.

It has already been suggested (cf. section 3.2) that the intervention of paragraph-marked (henceforth "p-marked") boundaries does have an observable impact upon referential choice, and allowing such break-points to act as heuristic guides has the additional advantage of enabling us to organize the available data into more approachable categories.[12] When, however, we approach the problem of characterizing the precise nature of this relationship, particularly in subject-switching contexts, it is important to note that, because in many cases several different factors appeared to contribute to the encoder's selection of a particular referential form, it proved extremely difficult, if not impossible, to separate and measure the various motivating forces in any truly objective fashion.[13] However, unlike Clancy (1980:170ff.), who chooses to discuss only cases of subject-preservation across episode-boundaries, I do not believe such analyti-

cal difficulties should be allowed to impede the attempt to describe some of the more general patterns involved, for some interesting correlations are apparent between the various NP categories considered and their distribution within and across p-marked segments. Figure 5.4 (cf. Appendix XI) now summarizes written narrative data on the formal scatter of undifferentiated NP vs. elliptical subject reference with respect to these two environments.

FIGURE 5.4. Written Narrative Percentages of Nominal and Elliptical Subjects Selected where Prior Reference is within or across a Paragraph-Marked Boundary.



Two generalizations are immediately possible:

(13) The use of subject-ellipsis is relatively unusual if the referent so coded is separated from its immediately preceding control by a p-marked break -- only 36 out of 229 (15.7%) recorded tokens fell into this category (cf. section 5.1), compared with 958 of 1046 (91.6%) same-paragraph tokens.[14]

(14) The crossing of a p-marked boundary, on the other hand, tends to precipitate a shift to some form of nominal coreference -- 84.3% (193/229) of the attested cases (cf. section 5.2).[15]

5.1. Subject-ellipsis within and across p-marked boundaries

The occurrence of the highly continuous subject-ellipsis strategy within the same p-marked segment as prior referential control is clearly the distributional norm -- 91.6% (958/1046) of attested cases -- and has been exemplified on so many occasions throughout this study that I shall say no more on the matter. It is the residual 15.7% (36/229) of elliptical cases where prior mention is located within a different paragraph which is somewhat unusual and so worthy of some comment. Of the 36 tokens recorded, the majority -- 24 (66.7%) -- occurred

where the prior mention was in fact the same subject argument of the final sentence of the immediately preceding paragraph;[16] the remaining 12 cases (33.3%) were found in slightly more disruptive discourse circumstances. In other words, even though subject-ellipsis is marginally permissible across p-marked boundaries, it tends not to be used unless it is part of a relatively integrated, same-subject sequence. Excerpts (15-17) illustrate this tendency.

(15) a. Ø Suka shirya,
SUBJ-Ø they-PFV get ready
b. Ø suka dunguma zuwa Bila.
SUBJ-Ø they-PFV troop off to Bila-LOC
c. Da Ø suka isa
when SUBJ-Ø they-PFV arrive
d. Ø suka nemi
SUBJ-Ø they-PFV seek
e. Ø su ga sarki...
SUBJ-Ø they-SUBJCTV see emir-DO

(Imam 1970:7)

'Ø (They) got ready and Ø trooped off to Bila.
When Ø (they) arrived Ø (they) sought an audience with the emir...'

The third person plural <u>suka</u> ellipsis strategy used at the beginning of the new p-marked boundary in (15c) follows upon a minor spatio-temporal shift introduced by the temporal conjunction <u>da</u> 'when', but is still an integral part of a chain of same-subject actions initiated in (15a). A number of the 24 tokens were encountered in comparable circumstances.[17]

More explicit and radical time-phrases can be interposed, but with the possible exception of marginal cases like (21) below, the same-subject constraint is generally observed, as illustrated in (16d):

(16) a. Ø Suka yi aurarraki,
SUBJ-Ø they-PFV do intermarriage
b. Ø suka zauna nan,
SUBJ-Ø they-PFV settle there-LOC
c. su da gari-n-su sai saƙo
they-IND PRO and town-of-them only message
ko yaushe wuri-n iyaye-n-su.
always place-of parents-of-them
d. Ran nan da dare Ø su-na
Day DEM at night-ADV SUBJ-Ø they-IMPFV
zaune da sarki...
sit-STAT with emir-ASSOC

(Imam 1970:9)

'Ø (They) intermarried and Ø settled there, and the only contact they had with their town was via messages to their parents.
One evening Ø (they) were sitting down with the emir...'

Example (17) exemplifies the same phenomenon:

(17) a. ...har ma Ø ya-kan kira karnuka ko
even and SUBJ-Ø he-HABIT call dogs or
kaji
chickens-DO
b. Ø ya ba su Ø.
SUBJ-Ø he-SUBJCTV give them-DO PRO DO-Ø
c. Ran goma sha biyu ga wannan wata,
Day ten and two of DEM month-ADV
watau dare-n Mauludi,
that is eve-of Mowlud
d. Ø ya-na barci
SUBJ-Ø he-IMPFV sleep
e. sai Ø ya yi mafarki...
then SUBJ-Ø he-PFV do dream

(Imam 1970:11)

'...Ø (he) would even call dogs and chickens and Ø give them (food).
On the twelfth day of that month, the eve of the Mowlud Festival, Ø (he) was sleeping, when Ø (he) had a dream...'

In (17d), the subject-ellipsis strategy yana (third person singular masculine imperfective) picks up the preceding (subjunctive) zero-strategy in (17b), despite the intervention of a quite "heavy" paragraph-initial temporal reorientation in (17c).

Fragment (18) is from one of the Pear Film accounts:

(18) a. ...sai wani yaro, sai ya ga hula-r.
then IS boy-SUBJ then he-PFV see cap-DET-DO
b. Sai Ø ya koma mishi
then SUBJ-Ø he-PFV return to him-IO PRO
da ita,
with it-IND PRO-ASSOC
c. Ø ya je
SUBJ-Ø he-PFV go
d. Ø ya ba shi Ø,
SUBJ-Ø he-PFV give him-DO PRO DO-Ø
e. Ø ya ce ga hula-r-sa.
SUBJ-Ø he-PFV say here is hat-of-his
f. Shikenan, da Ø ya ba
That was that, when SUBJ-Ø he-PFV give
shi hula-r...
him-DO PRO hat-DET-DO

(Speaker 1)

'...then a boy spotted the cap. Then Ø (he) returned it to him, Ø (he) went and Ø gave him (it) and Ø said here was his hat.
That was that, when Ø (he) had given him the hat...'

Sentence (18f) seems to coincide with an episode-boundary.[18] The speaker has been describing the details of the "helping" episode, and hesitates before proceeding to

relate the events of the "exchange" episode. Again, the absence of any participant switches, plus the fact that the speaker repeats some of the actions with a backtracking anterior-perfective in clause (18f), allows him to continue with the same-subject ellipsis strategy. Notice too the use of the connector shikenan 'that was that' at the beginning of (18f), a compound used typically to mark the onset of a new episode or event-structure in narrative tracts (cf. Jaggar 1982b).

Turning now to those paragraph-initial subject-ellipsis instances where immediate prior control is not coreferential, a breakdown of the 12 attested type-tokens reveals that these unusual cases were also subject to the following restrictive acceptability-conditions: 6 were examples of inclusive third person plural subject-ellipsis, i.e. the strategy simply gathered together several arguments which had been the subjects of the immediately preceding discourse; and in the remaining 6 cases, obligatory subject-agreement on the preverbal auxiliary was always sufficient to guarantee non-ambiguity of reference. Fragments (19-22) provide illustrations of both the above categories:

(19) a. Yaro ya ce,
boy-SUBJ he-PFV say
b. "Kurwa-ta kur, ka ci kan-n-ka,
soul-my bitter you-SUBJCTV eat head-of-you-DO
c. ka sha baki-n ruwa!"
you-SUBJCTV drink black-of water-DO

d. Aljani dai bai ce masa
jinn-SUBJ however NEG-he-PFV say to him-IO PRO
kome ba.
anything NEG
e. Ø Su-na nan zaune...
SUBJ-Ø they-IMPFV there-LOC sit-STATIVE

(Imam 1970:23)

'The boy said, "My soul is bitter (i.e. beware of devouring it), so eat your own and drink black water!" The jinn, however, said nothing to him. There Ø (they) were sitting down...'

The paragraph-initial imperfective subject-Ø strategy suna in (19e), with third-person plural agreement, though strictly speaking a different grammatical subject from the one in the prior, paragraph-final clause (19d) -- 'the jinn' -- could hardly be analyzed as an instance of discontinuous subject-switching -- it simply coalesces both preceding subject participants, 'the boy' (19a) and 'the jinn' (19d).

Excerpts (20-21) illustrate environments in which ellipsis can be exploited to encode genuine cases of cross-paragraph subject-switching, but only in circumstances where aux-agreement renders potential ambiguity impossible.

(20) a. ...amma fa ba a ce
but indeed NEG IMPERS-PFV say
b. in Ø ya je
when SUBJ-Ø he-PFV go
c. Ø ya ta da shi ba,
SUBJ-Ø he-SUBJNCTV awaken him-IND PRO NEG
d. ko motsi mai karfi ma kada kowa
even movement with loudness and NEG anyone-SUBJ

ya yi.
he-SUBJNCTV do

e. Ø Suka isa,
SUBJ-Ø they-PFV arrive

f. Ø ta tarad da daki-n rufe da
SUBJ-Ø she-PFV find hut-DET close-STAT with
asabari...
reed mat-ASSOC

(Imam 1970:25)

'...but it hadn't been said that when Ø (he/the grandfather) went (in) Ø (he) should awaken him (the grandson), nc-one should make so much as a sound.
Ø (They) arrived and Ø (she/the daughter) found the hut closed with a reed mat...'

The referent of the (perfective third person singular feminine) subject-zero strategy ta in (20f) is 'the daughter', last mentioned in the previous paragraph, with 'the grandfather' intervening as a subject-switch argument in (20b-c). However, because 'the daughter' is the only female participant in the immediate vicinity, use of the subject-ellipsis strategy poses no problems for correct interpretation -- cf. fragment (37) below, however. And in (21):

(21) a. Ko da gari ya waye
As soon as dawn-SUBJ it-PFV break
sai Ø suka shiga shawara-r
then SUBJ-Ø they-PFV enter deciding-of
b. wanda za-su wa sabo-n sarki.
one who-REL-IO FUT-they to new-of emir

(Imam 1970:38)

'At daybreak Ø (they) began to decide whom Ø (they) should appoint the new king (of the thieves).'

use of subject-Ø device suka in the paragraph-initial (21a), despite the presence in this clause of a strongly reorientational time-phrase, is permissible because the prior context is only marginally disruptive/discontinuous -- the intervening material includes only a negative, nondeployable subject argument 'no-one', and a paragraph-final evaluative clause.

Extract (22) provides a similar example from a Pear Film account:

(22) a. Da Ø ya tsinko wadannan 'ya-n
when SUBJ-Ø he-PFV pluck DEM DIMIN-of
itatuwa
trees-DO
b. sai Ø ya sanya Ø ciki-n kwando.
then SUBJ-Ø he-PFV put DO-Ø inside-of basket-LOC
c. To, sai can, sai da an jera
OK then later until IMPERS-PFV line up
kwanduna guda uku,
baskets unit three-DO
d. sai Ø ya-na nan,
then SUBJ-Ø he-IMPFV there-LOC
e. Ø ya-na ta tsinka Ø...
SUBJ-Ø he-IMPFV keep on pluck-VN DO-Ø

(Speaker 2)

'When Ø (he/the mango man) had picked those fruits Ø (he) put (them) in a basket.
OK, later on, when three baskets had been lined up, there Ø (he) was, picking (the mangoes)...'

Between the two cross-paragraph elliptical subject references to 'the mango man' in (22b) and (22d), we have the intervention of a temporally-signaled transition from the "basket-filling" episode to a descriptive background segment in (22c), and thence to a new event-structure initi-

ated in (22d), in addition to a switch to the (perfective) impersonal auxiliary an in (22c). Again, I would resist any attempt to analyze this as a genuine case of subject-switching. The referentially vacuous, non-specific impersonal auxiliary is part of a fairly coherent event-chain, with the inference that 'the mango man' is still the agential force in (22c).

Finally on the matter of subject-ellipsis, fragment (23) nicely illustrates two cross-paragraph cases:

(23) a. ∅ Ta ci gaba da suya,
SUBJ-∅ she-IMPFV continue with cooking-VN
b. ƙanshi ∅-na jifa-r-ta,
aroma-SUBJ CONC-∅-IMPFV throw-of-her
...intervening clauses...
c. Kafin ∅ su soyu
before SUBJ-∅ they-SUBJCTV fry-PASS
d. ∅ ta kusa cinye rabi-n kaza.
SUBJ-∅ she-PFV approach eat up half-of chicken-DO

(Imam 1970:16)

'∅ (She) carried on with the cooking, the aroma overpowering her...Before ∅ (they/the chickens) were fully fried, ∅ (she) had almost eaten uphalf a chicken.'

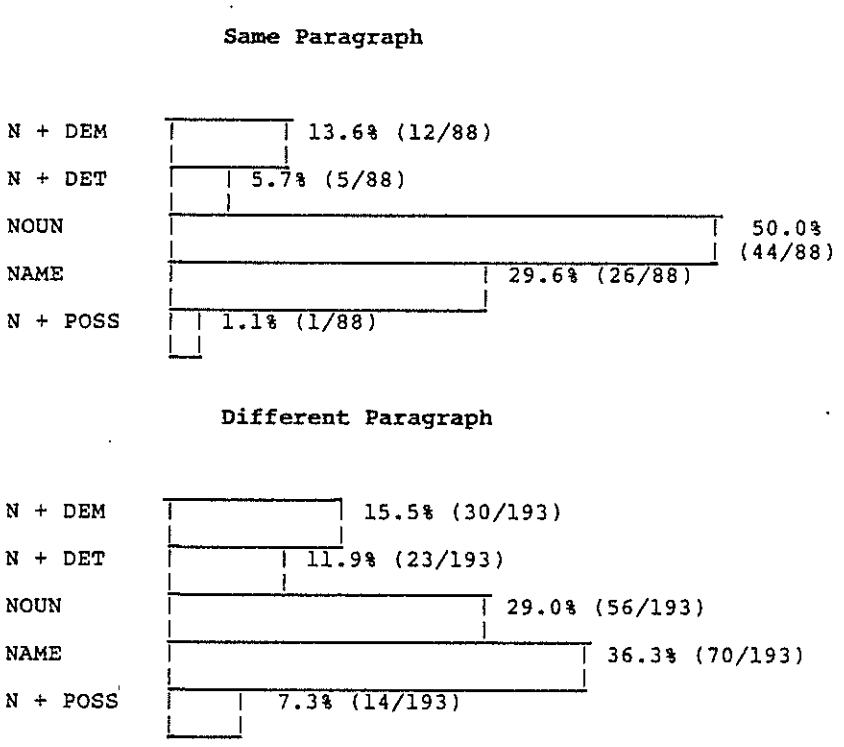
The paragraph-initial (third person feminine singular perfective) elliptical form ta in (23a) is part of a same-subject sequence, and since its referent is the only female participant in the plot, unique identification is assured. And use of the similarly nonovert (third person plural subjunctive) elliptical device su 'they (the chickens)'in (23c), despite the fact that its prior control is

in the preceding paragraph, is acceptable for two reasons: it is the only plural argument in the vicinity; and the lexicosemantics of the following passive verb soyu 'be fully fried' (cf. Jaggar 1982a) guarantee the identity of the ellipted subject.

5.2. Subject nominal coreference within and across paragraph-marked boundaries

Figure 5.5 (cf. Appendix XII) presents a detailed breakdown of the 281 general NP tokens in Figure 5.4, and summarizes data on the distribution of all five NP combinations with respect to whether the immediately preceding discourse reference was located within the same paragraph or in a preceding paragraph.

FIGURE 5.5. **Written Narrative Percentages of Differential NP Configurations Selected where Prior Mention is within and across Paragraph-Marked Boundaries.**



The data in Figure 5.5 allow a few immediate comments:

(24) Excepting the N + possessor class, which, as we have noted before, constitutes, along with proper names, a rather special referential category -- the deictic-bearing N + demonstrative (30/193 or 15.5%) and N + determiner (23/193 or 11.9%) groups have, between them, almost as many different-paragraph tokens (53/193 or 27.4%) as the class of simple nouns (56/193 or 29.0%). Inspection of the corresponding same-paragraph values, on the other hand, reveals that the same two N + demonstrative (12/88 or 13.6%) and N + determiner (5/88 or 5.7%) classes together account for a substantially lower proportion (17/88 or 19.3%) than the category of simple nominals (44/88 or 50.0%).[19]

A glance back at Figure 5.1(a) will remind the reader that only 5 (3.6%) out of a total 137 same subject/new sentence examples in the written texts were NP-coded, and in section 3.2 it was pointed out that several of these cases coincided with paragraph-marked boundaries. Before moving on to consideration of the normative instances of NP-signaled subject-switching, therefore, it is instructive to illustrate and discuss these rarer types. Extracts (25-27) in fact exhaust the category in question, with fragment (25) instantiating a demonstrative-marked same-subject NP:

(25) a. Sarki ya ce,
emir-SUBJ he-PFV say
b. "I, ai ko lalle wannan gardama
yes well and certainly DEM quarrel-SUBJ

ta isa
it-PFV be enough
c. a rabo da gida,
IMPERS-SUBJCTV part with home-ASSOC
d. a shiga duniya nema-n inda
IMPERS-SUBJCTV enter world seeking-of-VN where
za-a raba ta."
FUT-IMPERS resolve it-DO PRO
e. To, sarki-n nan kuwa Ø-na
Now emir-DET DEM-SUBJ and CONC-Ø-IMPFV
kiwo-n balbelu kama-r dari.
tending-of egrets like-of hundred

(Imam 1970:7)

'The emir said, "Yes, this quarrel is serious enough for one to travel afar seeking to have it resolved."
Now this emir kept about a hundred cattle egrets.'

In (25a), the bare nominal sarki 'the emir' is the subject of the performative verb 'say', and governs the quoted direct speech in clauses (25b-d) -- cf. section 5.3. The same referent then appears as a demonstrative-marked NP **sarkin nan** 'this emir' at the onset of the new paragraph in (25e), and I would maintain that the writer selects this escalated coding device in order to alert his reader to the fact that a radical transition is about to take place -- (25e) in fact marks the beginning of a major new episode, one which provides the critical background setting and motivation for the rest of the plot.[20] Russell Schuh (p.c.) has also suggested that use of a demonstrative in such contexts, as opposed to the determiner, might convey a contrastive interpretation, in which case a more accurate gloss of (25e) would be something like 'now this

particular emir (unlike others)...'.

Consider now extract (26):

(26) a. Kaka-n ya yi murmushi,
grandfather-DET-SUBJ he-PFV do smile
b. Ø ya koma gida.
SUBJ-Ø he-PFV return home-LOC
c. Gari Ø-na waye-wa
dawn-SUBJ CONC-Ø-IMPFV break-VN
d. sai **kaka-n** ya tafi
then **grandfather-DET-SUBJ** he-PFV go
e. Ø ya gaya wa sarki...
SUBJ-Ø he-PFV tell to emir-IO

(Imam 1970:21)

'The grandfather smiled and Ø returned home.
At daybreak **the grandfather** went and Ø told the emir...'

Assuming, in the general case, a strategy of same-subject interpretation for sequences of ellipted subjects, the putative use of ellipsis in (26d) would have been quite unambiguous. The occurrence, instead, of the determiner-coded subject N **kakan** 'the grandfather', is activated by several factors, including: the marked shift in temporal setting, signaled by the paragraph-initial time phrase 'at daybreak' in (26c);[21] and a change in both location and event-structure, initiated by 'the grandfather's' visit to 'the emir'.

An additional, language-specific factor is at work in fragment (26) which must be taken into account. There is an observable reluctance to use a bare nominal to denote such specific relational categories as 'father, mother,

grandfather, grandmother, husband, wife' etc. -- such inalienable referents require modification of some kind (cf. Jaggar 1981:54, 55n). Of the three other explicit options available, moreover, selection of a "high-profile" demonstrative in every instance would clearly be unacceptable, violating Grice's (1975:45) maxim of quantity, which leaves the determiner or following possessor argument as the only remaining referential candidates. The particular story in which 'the grandfather' referent occurs -- (Imam 1970:20-26) -- contains a total of 13 explicit mentions of the character, with the category-by-category breakdown as follows: 7 N + possessor; 5 N + determiner (including 26a, d); 1 N + demonstrative.[22] We shall encounter additional examples of this coding constraint when we come to consider the referential strategies exploited in turn-taking dialogue in the final section 5.3.

Finally, consider fragment (27):

(27) a. Da zuwa-n baƙauye
with coming-of-PART villager
b. aka yi ciniki-n ice,
IMPERS-PFV do bargaining-of wood
c. Ø ya sallama,
SUBJ-Ø he-PFV agree
d. aka ce Ø ya sauke Ø
IMPERS-PFV say SUBJ-Ø he-SUBJCTV put down DO-Ø
e. Ø ya shiga da shi ciki-n
SUBJ-Ø he-SUBJCTV enter with it-PRO inside-of
barga.
stable-LOC
f. **Baƙauye** Ø-na kunna kai
villager-SUBJ CONC-Ø-IMPFV poke head
zaure na biyu,
entrance-hut of two-LOC

g. sai Ø ya ji
then SUBJ-Ø he-PFV feel
h. ƙanshi-n mai ya buge shi.
aroma-of oil-SUBJ it-PFV hit him-DO PRO

(Imam 1970:30)

'As soon as the villager arrived they ('one') bargained over the wood, Ø (he) agreed to a price and they said Ø (he) should put (it) down and take it into the stables.
The villager was poking his head into the second entrance-hut when Ø (he) felt the aroma of oil hit him.'

In (27f), occurrence of the simple noun subject baƙauye 'the villager' at the start of the new paragraph, despite the fact that the same referent was the grammatical subject of the immediately preceding clause (27e), is arguably attributable to the switch in event-structure or action from the "bargaining" episode described in clauses (27b-e) to the new "food" episode, in addition to a locational shift.

The data in examples (25-27) have several implications. Notice, firstly, that they are all damaging to Givón's (1981:4) quite specific claim that a paragraph-initial "definite or returning topic may return after a large gap of absence, a smaller one or a smaller one yet (though still not recurring in successive clauses)" [my emphasis] -- the paragraph-initial "topic" NP baƙauye 'the villager' in (27f), for instance, does recur in a clause successive to its prior appearance. Secondly, they illustrate the manner in which p-marked discourse breaks can

act to precipitate selection of a more explicit coreferential device, even in the less disruptive environment of subject-maintenance. At the same time, of course, it is important to point out the range of referential forms used -- N + demonstrative in (25e), N + determiner in (26d), and simple noun in (27f) -- a fact which serves to underscore the difficulty of formulating any hard-and-fast "rules" to account for the formal distribution of all the ongoing choices attested.

The above reservations notwithstanding, the data seem to suggest a general pattern which is basically consonant with the regularities predicted by the deixis:marking hypothesis, and which may be informally stated as follows: the larger the number of discourse-related factors clustering at a given paragraph-signaled break and so adding to the reorientation burden on both the encoder and decoder, the greater the tendency to exploit a high-density coding device. In other words, the degree of processing difficulty involved in transferring from one focus of attention to another will increase with the amount of reorientation entailed. The more burdensome this realignment, the more inclined the encoder will be to select a higher deictic option, either to assist himself/herself in the cognitive task of verbalizing the transition, or to facilitate the decoder's processing of the switch. The implied correlations could not, of course, be absolute, simply because

the precise intersection of the above factors, and their ultimate impact on referential choice, are given to some variation, and so we are limited to the generalization that, all things equal, there will be discourse circumstances in which the occurrence (or not) of a given referential strategy will either be likely, optional, or unusual. Furthermore, since a complete survey of all the attested cases is obviously beyond the scope of this chapter, I shall limit myself to illustrating the observed regularities with some of the more stereotypical category-tokens, concentrating, as ever, on the three configurations of noun + demonstrative, noun + determiner, and simple noun.

Within the functional domain of NP-coded subject-switching, extracts (28-31) typify the kinds of maximally discontinuous p-marked boundaries which regularly elicit the use of a demonstrative to reintroduce a switch-subject referent.

(28) a. ...Narimi Ø-na zulumi-n abi-n,
Narimi-SUBJ CONC-Ø-IMPFV anxious-of thing-DET
b. Ø ya-na ce-wa,
SUBJ-Ø he-IMPFV say-VN
c. "Lalle baki dai shi ya-ke yanka
surely mouth-SUBJ well IND PRO it-IMPFV cut
wuya."
throat-DO
d. Magariba Ø-na yi
evening-SUBJ CONC-Ø-IMPFV do-VN
sai yaro-n nan ya ɗauki 'ya-r
then boy-DET DEM-SUBJ he-PFV take DIMIN-of
tabarma-r-sa
mat-of-his-DO

e. Ø SUBJ-Ø he-PFV ya go tafi foot-of gindi-n tamarind tree-DET tsamiya-n DEM-LOC nan...

(Imam 1970:23)

'...Narimi (the grandfather) was anxious about the matter, Ø saying, "It's the mouth that cuts the throat" (i.e. be careful what you say).
As evening arrived **that boy** (the grandson) picked up his little mat and Ø went to the foot of that tamarind tree...'

The demonstrative-bearing subject NP **yaron nan** 'that boy' in (28d), last mentioned in the preceding paragraph, coincides with a strong p-marked switch in subject, spatio-temporal setting, and action-structure. Similarly, in (29):

(29) a. Da when rana sun-SUBJ ta it-PFV fito come out
b. Ø SUBJ-Ø he-PFV ya take dauki DIMIN-of 'ya-r mat-of-his-DO tabarma-r-sa,
c. Ø SUBJ-Ø he-PFV ya head nufi for gida-n palace-of sarki, emir
d. Ø SUBJ-Ø he-PFV ya say ce
e. Ø SUBJ-Ø he-PFV ya return dawo.
f. Mutane people-SUBJ za-su FUT-they yi do musu-n contradict-of
g. Ø SUBJ-Ø bai NEG-he-PFV kwana spend night ba, NEG
h. sai then **malama-n teachers-DET** **nan hudu DEM four-SUBJ** suka they-PFV shaida confirm
i. Ø SUBJ-Ø he-PFV ya spend kwana. night

(Imam 1970:24)

'When the sun got up Ø (he) picked up his little mat, Ø headed for the emir's palace and said Ø (he) was back. The people were about to deny that Ø (he) had spent the night (amidst the spirits)

when **those four teachers** confirmed that Ø (he) had spent the night (there).'

the occurrence of a new episode, along with a temporal reorientation, elicits use of the demonstrative-marked switch-subject **malaman nan (hudu)** 'those (four) teachers' in (29h). In this case too, the fact that over 40 separate clauses and 5 p-marked boundaries separate this same NP from its prior mention, surely acts as an additional force in persuading the writer to use the high-deixis demonstrative (cf. Clancy's (1980:173ff.) similar findings for English and Japanese). And Russell Schuh (p.c.) has reminded me that that the presence in (29f) of a maximally similar third person plural human subject <u>mutane</u> 'the people' is surely also pertinent to the choice encountered in (29h). Notice too the use of subject-ellipsis in (29b), coinciding with the onset of a new paragraph in (29a), and following closely upon a same subject reference in the final clause of the preceding paragraph -- cf. section 5.1. And in (30b):

(30) a. Kowa everyone-SUBJ ya he-PFV ce say shi he-IND PRO za-a FUT-IMPERS ba. give
b. Daga nan sai then **tsohuwa-n old woman-PRM** **nan DEM-SUBJ** ta she-PFV ce... say

(Imam 1970:38)

'Everyone said that he should be given (the position of "King of the Thieves"). Then **that old woman** said...'

prior reference to the demonstrative-coded switch-subject **tsohuwan nan** 'that old woman' spans several p-marked segments and a total of 90 or so clauses. It is worth pointing out too that since this same referent was the only female participant in the story concerned, use of subject-ellipsis (with third person feminine agreement) would have been completely unambiguous. However, the passage of several episode boundaries appears to preclude this possibility and serves to elicit nominal coreference.

In the two Pear Film accounts, there was one point where both speakers felt a need to use a demonstrative -- where 'the mango man' is reintroduced into the plot after a lengthy absence and following several important twists and turns in the story. Extract (31) is illustrative:

(31) a. Ø Ya-na tafiya,
SUBJ-Ø he-IMPFV go-VN
b. Ø ya-na tafiya,
SUBJ-Ø he-IMPFV go-VN
c. can sai **mutumi-n can** ya juya...
later then **man-PRM DEM-SUBJ** he-PFV turn around

(Speaker 2)

'Ø (He/the boy) was going along when that man (the mango-man) turned round...'

In (31c) the speaker uses a <u>can</u>-coded N + demonstrative **mutumin can** to reintroduce 'the mango man' at a point of subject-switching.[23]

Finally, fragment (32) exemplifies the use of a demonstrative -- this time in a more continuous environment.

The demonstrative in question is used in a Pear Film account to provide a referent -- 'the boys' -- with a "send-off":

(32) a. Sai Ø ya ga
then SUBJ-Ø he-PFV see
b. yara kawai sun fice,
boys-SUBJ just they-PFV pass by
c. Ø su-na sha-n mangwaro.
SUBJ-Ø they-IMPFV drink-of-VN mangoes
d. Sai Ø ya tsaya
then SUBJ-Ø he-PFV stop
e. Ø ya-na tunani-n
SUBJ-Ø he-IMPFV wonder-of
f. yaya aka yi
how IMPERS-PFV do
g. **yara-n nan suka** sami mangwaro-n.
boys-DET DEM-SUBJ they-PFV get mangoes-DET-DO
h. Karshe-n labari kenan.
end-of story COP

(Speaker 2)

'Then Ø (he) saw the boys had just passed by Ø eating mangoes. Then Ø (he) stopped and Ø was wondering how it was that those boys had got the mangoes. And that's the end of the story.'

In (32g), the speaker signs off with a full demonstrative on the subject NP **yaran nan** 'those boys', despite the fact that the same referent was last mentioned only 4 clauses earlier, and no potentially confusing, maximally similar arguments have intervened.[24]

Moving on to the N + determiner combination, as already noted, the formal distribution of this category overlaps partially with that of demonstrative-marked nominals in that it is frequently encountered in relatively discontinuous, cross-paragraph environments, underscoring

the relative freedom allowed in the selection of referential strategies at such boundaries. Fragment (33) illustrates the use of this particular strategy to specify a switch to a subject whose prior referential appearance is in the preceding paragraph:

(33) a. Ø Su-na nan
SUBJ-Ø they-IMPFV there-LOC
b. Ø su-na shawara-r gudu ke nan,
SUBJ-Ø they-IMPFV contemplate-of fleeing-VN COP
c. sai attajiri-n ya sake
then merchant-DET-SUBJ he-PFV do again
yi-n kira.
do-of-VN call-VN

(Imam 1970:32)

'There Ø (they) were, contemplating running away, when **the merchant** summoned (a servant) again.'

Occurrence of the determiner-equipped new subject NP **attajirin** 'the merchant' in (33c) coincides with the shift to a fresh event-structure, although no spatio-temporal reorientation is involved in this instance.

Extract (34) instantiates the use of a determiner to specify a switch-subject referent -- **yaron** 'the boy' (part of a conjoined subject NP in fact) -- following a paragraph-initial subordinate clause which shifts the temporal scenario:

(34) a. Da aka yi salla-r azahar,
when IMPERS-PFV do prayer-of mid-afternoon
b. da **yaro-n** da kaka-n-sa
and **boy-DET-SUBJ** and grandfather-of-his-SUBJ
suka tafi gida-n sarki.
they-PFV go palace-of emir-LOC

(Imam 1970:22)

'When the mid-afternoon prayer had taken place, **the boy** and his grandfather went to the emir's palace.'

Turning finally to brief consideration of the simple noun category, Figure 5.5 shows that of the 100 tokens in the sample, 44.0% (44/100) fell within the same paragraph as prior reference, with 56.0% (56/100) occurring in a different paragraph. The same/different paragraph counts are thus closer than for any other category, a finding which serves to support earlier claims that low deixis-signaling bare N's tend to be utilized in less disruptive discourse contexts. Several examples of the role of simple N's in straightforward subject-switch contexts have already been provided in section 4.2 above, and one more illustration should suffice. Fragment (35) exemplifies a sequence disrupted only by shifts in subject, and in which simple nouns are exploited to encode these switches:

(35) a. **Sarki-n fawa** ya sa su
chief-of butchering-SUBJ he-PFV put them-DO PRO
gaba har majalisa-r alƙali.
in front-LOC until chambers-of judge.
b. **Alƙali** ya tambayi mai kudi-n
judge-SUBJ he-PFV question one with money-DET
nan,
DEM-DO
c. Ø ya bayyana masa yadda aka
SUBJ-Ø he-PFV explain to him-IO how IMPERS-PFV
yi.
do
d. Da **alƙali** ya ji haka...
when **judge-SUBJ** he-PFV hear this

(Imam 1970:27)

'**The chief butcher** marched them right to the judge's chambers. **The judge** questioned the owner of the money and Ø (he) explained to him what had happened. When **the judge** heard this...'

5.3. Direct speech quotations and referential choice

Since the written narratives examined were regularly interspersed with quoted direct speech, it seems appropriate to conclude this study with some brief remarks about the effect of this type of intervening material on referential choice. The data derive from a count of the 130 subject arguments immediately following the closing quotation boundaries of embedded direct speech segments in Imam (1970:6-42). Of these 130 type-tokens, 86 (66.2%) were instances of NP's used in subject-switching contexts, and the occurrence of some form of NP in this position clearly represents the norm -- cf. section 4. Subject-ellipsis accounts for the remaining 44 (33.8%) cases. Basically, the same generalizations expressed in sections 5.1 and 5.2 about the subject-Ø:NP choices across p-marked boundaries extend to the domain now under consideration.[25] The evidence at hand suggests that the interposing of direct speech is often considered enough of a break in the narrative discourse flow to favor exploitation of a following explicit NP strategy in the subject slot, i.e., even in

those contexts where use of nonovert subject-ellipsis would not have produced referent ambiguity.

With respect to the cases of subject-ellipsis encountered in this environment, the occurrence of this strategy is subject to the kinds of constraints we have noted elsewhere, i.e. it can be used without resultant ambiguity either because obligatory subject-agreement on the preverbal auxiliary supplies unique identification in cases of subject/speaker-switching, or because the elliptical device in fact represents a preserved subject. In more discontinuous and disruptive contexts, the overwhelming tendency is to use some form of NP. The following illustrative fragments (36-38) all contain items exemplifying both lexical NP subjects and ellipted subjects.

(36) a. **Mutane** suka ce,
people-SUBJ they-PFV say
b. "I, Allah ya gafarta malam."
yes Allah-SUBJ he-SUBJCTV forgive teacher
c. Ø Ya dubi Filani,
SUBJ-Ø he-PFV look at Fulani-DO
d. Ø ya ce,
SUBJ-Ø he-PFV say
e. "Ku kuma kun gani hakanan ne ko?"
you-IND PRO and you-PFV see thus COP or what
f. **Filani-n** suka ce,
Fulani-DET-SUBJ they-PFV say
g. "I, Allah shi dadè da ra-n-ka."
yes Allah-SUBJ he-SUBJCTV prolong life-of-you
h. **Alƙali** ya ce...
judge-SUBJ he-PFV say

(Imam 1970:28)

'**The people** said, "Yes, (may Allah forgive and bless the) teacher."
He looked at the Fulani men and said, "And you,

do you agree to this?"
The Fulani said, "Yes, sir."
The judge said...'

Observe that the only instance in this turn-taking exchange of a switch subject-speaker not coded by some form of lexical NP occurs in (36c-d), where ellipsis is used to refer to 'the judge'. In such circumstances, the obligatory auxiliary-agreement is sufficient to distinguish this singular argument (also referred to in the vocative in (36b)) from the remaining two maximally similar plural arguments in the piece -- **mutane** 'the people' (36a) and **Filanin** 'the Fulani men' (36f). These two referents are both specified with some form of NP. Notice, however, that in (36h) Imam reverts immediately to the more usual nominal-coded speaker-subject switching for the same referent 'the judge', despite the fact that the same conditions appear to hold. It is cases such as this which exemplify the preference for post-direct speech overt mention.[26] Extract (37) provides additional illustration:

(37) a. Kalalatu kuma ta ruga waje-n Kalala,
Kalalatu-SUBJ and she-PFV rush place-of Kalala
b. Ø ta ce,
SUBJ-Ø she-PFV say
"direct speech fragment"
c. Kalala ya ce,
Kalala-SUBJ he-PFV say
"direct speech fragment"
d. Kalalatu ta ce,
Kalalatu-SUBJ she-PFV say
"direct speech fragment"

e. Kalala ya fusata,
Kalala-SUBJ he-PFV become angry
f. Ø ya ce...
SUBJ-Ø he-PFV say

(Imam 1970:18)

'Kalalatu rushed to Kalala and Ø said, " "
Kalala said, " "
Kalalatu said, " "
Kalala became angry and said...'

All three references to the turn-taking subjects in (37c-e) are coded with full proper names, despite the additional fact that Kalalatu is the only female character in the story and so could have been ellipted without any resultant ambiguity.

A substantial proportion of immediate post-direct speech tokens were encountered where the ellipted subject argument was in fact coreferential with the subject-speaker controlling the preceding direct speech fragment in question. Extract (38) is illustrative:

(38) a. Sarki ya tambaye shi.
emir-SUBJ he-PFV question him-DO PRO
b. Isa ya ce,
Isa-SUBJ he-PFV say
"direct speech segment"
c. Ø Ya cire rawani,
SUBJ-Ø he-PFV take off turban-DO
d. Ø ya ce,
SUBJ-Ø he-PFV say
"direct speech segment"
e. Shu'aibu, da-n wanzam, ya dubi
Shu'aibu son-of barber-SUBJ he-PFV look at
Isa Lamiri,
Isa Lamiri-DO
f. Ø ya ce...
SUBJ-Ø he-PFV say

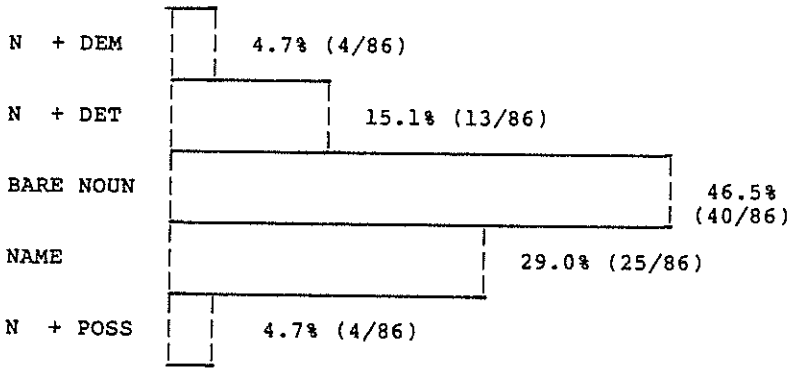
(Imam 1970:13)

'The emir questioned him. Isa said, "....." ∅ (He/Isa) pulled off his turban and ∅ said, "....." Shu'aibu the barber's son looked at Isa Lamiri and said...'

Clauses (38b-c) illustrate the relatively continuous environment in which the selection of ellipsis is acceptable despite the intervention of direct speech -- the preverbal auxiliary ya in (38c) simply picks up the proper name subject Isa in (38b). Same-subject interpretation is favored, and in addition to this, the topic of the preceding direct speech clauses concerns Isa's claim that Shu'aibu the barber has not given him a proper haircut -- Isa has to take off his turban to prove his point. Reference to the remaining three maximally similar interactants, on the other hand, is achieved via nominal mention of some kind, in order to avoid potential confusion.

As noted earlier, following direct speech, overt lexical reference to a switch subject is the most frequently-encountered option -- 66.2% (86/130) -- especially in turn-taking conversational situations where the putative use of an inexplicit zero strategy might otherwise result in potential ambiguity. Consistent with the discriminant approach which has provided the baseline for the last two chapters, Figure 5.6 (cf. Appendix XIII) now presents a category-based breakdown of these 86 general NP tokens.

Figure 5.6. Written Narrative Percentages of Differential NP Configurations Selected to Encode Subject Arguments immediately following "Direct Speech" Quotations.



The data in Figure 5.6 show that simple nouns (46.5%) and proper names (29.0%) represent the most commonly-selected options, and we have already noted some typical examples of both these categories. Concerning the two deictic-marked NP combinations, their post-direct speech distribution merits brief consideration, since it parallels the distribution of these same items as outlined in section 5.2.

Only 4 (4.7%) tokens of the "high-profile" N + demonstrative category occurred as switch-subjects following

quotation boundaries, and all but one of these cases entailed the kinds of additional disruptive factors which have been noted to consistently elicit a high-deixis demonstrative, i.e., a substantial gap to prior discourse mention, a radical shift in the action/event-sequence, spatiotemporal setting etc. -- cf. examples (28-31). I found only one example in a canonical turn-taking exchange:

(39) a. Mai gida ya ce,
compound head-SUBJ he-PFV say
b. "Ni? Duba na-ke yi."
Me-IND PRO Fortune-telling I-IMPFV do-VN
c. Baƙauye-n nan ya ce,
villager-DET DEM-SUBJ he-PFV say
d. "Ni in i-na so in koyi
Me if I-IMPFV want-VN I-SUBJCTV learn
duba,
fortune-telling-DO
e. na iya?"
I-FUT be able
f. Mai gida ya ce,
compound head-SUBJ he-PFV say
g. "Ka iya mana, ai ba
you-FUT be able of course well NEG EXIST
wuya."
difficulty
h. Baƙauye ya ce...
villager-SUBJ he-PFV say

(Imam 1970:31)

'The compound head said, "Me? I do fortune-telling."
That villager said, "Myself, if I wanted to learn fortune-telling, would I be able to?"
The compound head said, "Of course you'll be able to, it's not difficult."
The villager said...'

It is the switch-subject N + demonstrative baƙauyen nan

'that villager' in (39c) which is somewhat unusual, for it was the subject of the direct speech in the paragraph immediately preceding (39a). Notice, however, the immediate return in (39h) to the form usually exploited in such an environment -- the simple nominal baƙauye.

With regard to the determiner-marked cases, recall the observation made in section 5.2 that the determiner was regularly selected to modify inalienable kin-terms, in order to avoid use of a bare nominal -- cf. example (26d). Fragment (40) illustrates this same marking convention in the context of a conversational exchange:

(40) a. Uba-n ya ce, " "
father-DET-SUBJ he-PFV say
b. Yarinya ta ce, " "
girl-SUBJ she-PFV say
c. Uba-n ya yi dariya
father-DET-SUBJ he-PFV do laughter
d. Ø ya ce, " "
SUBJ-Ø he-PFV say
e. Yarinya-r ta ce, " "
girl-DET-SUBJ she-PFV say
f. Uwa-r ta yi dariya
mother-DET-SUBJ she-PFV do laughter
Ø ta ce, " "
SUBJ-Ø she-PFV say
g. Uba-n ya ce, " "
father-DET-SUBJ he-PFV say

(Imam 1970:24-25)

'The father said, " "
The girl said, " "
The father laughed and Ø said, " "
The girl said, " "
The mother laughed and said, " "
The father said, " "'

With the sole exception of the bare N yarinya 'the girl'

in (40b), all the subject-speaker kin-terms in this fragment are determiner-marked.

Outside the rather specialized domain of kin-terms, examples of determiner-coded switch-subject NP's in immediate post-direct speech contexts are available, but the precise contexts in which they occur tend not to be as disruptive as those which regularly elicit a demonstrative. Along with example (36f), therefore, consider the following:

(41) a. ...sai kaka-n ya yad da
then grandfather-DET-SUBJ he-PFV throw away
kara-n,
stalk-DET-DO
b. Ø ya ce,
SUBJ-Ø he-PFV say
c. "Kai, kai, kai, ni ne kaka-n-ka
hey hey hey me-IND PRO COP grandfather-of-you
Narimi,
Narimi
d. kada ka buge ni."
NEG you-SUBJCTV beat me-DO PRO
e. **Yaro-n** ya yi murmushi
boy-DET-SUBJ he-PFV do smile
f. Ø ya ce,
SUBJ-Ø he-PFV say
"direct speech fragment"
g. Narimi ya ce,
Narimi-SUBJ he-PFV say
"direct speech fragment"
h. **Yaro** ya ce...
boy-SUBJ he-PFV say

(Imam 1970:21)

'... then the grandfather threw away the stalk and
Ø said, "Hey, hey, hey, it's me your grandfather
Narimi, don't beat me."
The boy smiled and Ø said, " "
Narimi said, " "
The boy said, " "'

Notice the differential coding of the same subject referent in seemingly identical circumstances -- N + determiner **yaron** 'the boy' in (41e), compared with the more usual simple N **yaro** in (41h).

6. Summary

In the earlier part of this chapter I suggested that subject-switching appears to be a major factor behind the partitioning of narrative discourse into distinct sentences, and also in triggering an escalation from nonovert to overt referential coding. Attention was then focused on the impact of more radical discourse breaks on referential choice. It was proposed that switches in orientation are present in varying degrees at different junctures in discourse, and I pointed to the possibility of some general correlations with selection of referential forms. As Chafe (1980a:45) has observed, "First, we have identified various components of an orientation, not just one, and one or two or all of these components may be present at any particular point of transition. Changes of space, time, people, etc., tend to cluster, but they need not all be present at the same point. Beyond that, the components themselves are scalar: there may be more or less of a change in location, more or less of a shift to a new time frame, more or less of a change in protagonists, more or

less of a shift in background activity."

Finally, although it is clear that some of the quantified differences -- especially those characterizing the distributions of the categories noun, noun + demonstrative, and noun + determiner -- are not significant enough to permit any firm conclusions, I believe that some of the documented patterns do suggest some interesting cross-text regularities. I hope too, that this particular line of investigation, with little tradition behind it, may provide some helpful leads for similar studies in the future.

Notes to Chapter 5

[1]My approach to the data is based partially upon the analytical methodology followed by Clancy (1980:160ff.) in her statistically-based investigation of comparable phenomena in the English and Japanese "Pear Film" accounts, and I shall make regular reference to her study for purposes of comparison.

[2]An additional construction is in fact available for subject arguments, entailing the use of a stressed independent pronoun in position before the preverbal auxiliary, e.g.,

(1) a. ...Ø ya daura wani mutum-mutumi dogo,
SUBJ-Ø he-PFV tie IS effigy tall-DO
b. Ø ya sa masa fara-r riga,
SUBJ-Ø he-PFV put to him-IO PRO white-of gown-DO
c. **shi** kuma ya sa wa ka-n-sa
IND PRO-SUBJ-TOPIC and he-PFV put to head-of-his-IO
fara.
white earth-DO

(Imam 1970:20)

'...he (the grandfather) put together a tall effigy and put a white gown on it, and **he** put some white earth on himself.'

This pronominal strategy has not been included in the present analysis, however, mainly because its occurrence is normally restricted to contexts involving either topicalization or focus (cf. McConvell 1973; Schachter 1973; Jaggar 1978). In (1c), for example, the third person singular masculine independent pronoun **shi** 'he' is the topicalized subject of its clause. In addition, only a handful of such examples were in fact recorded -- 12 tokens in both the written and oral corpora.

[3]In the spontaneous oral narratives I transcribed, sentences, although not always clearcut, tended to be characterized by intonational downdrift, syntactic closure of some kind, and were separated by hesitations of approximately one second (cf. Chafe 1979b). See Chafe (1979b:162ff., 1980a) and Clancy (1980) for evidence supporting the view that sentences constitute processing units which are crucial both to the encoder's task of retrieving and packaging cognitive material, and to the

decoder's job of assimilating and storing information.

[4]Notice that I am including, within the subject category, arguments which are the grammatical subjects of complement object clauses, e.g. the (third person plural subjunctive) subject-Ø strategy su in (4e).

[5]Since the referent in question -- 'the Tuareg' -- is relatively new to the plot, having been introduced for the first time some 11 clauses earlier, the use of a modifying determiner in (6e) might perhaps be considered a marginal example of new referent-anchoring (cf. Chapter 4). However, as is so often the case when one is investigating the incidence of discourse-related linguistic expressions, it is difficult to tease apart the various conditioning factors. At the same time, it is worth noting that no examples of the heavyweight demonstrative strategy were encountered in the type of environment illustrated in (6).

[6]Li and Thompson (1979) document cases in Mandarin Chinese where the intervention of such contrastive elements precipitates an escalation from zero-anaphora to pronominalization. Notice that, along with some of the English forms recorded in Clancy (1980:179-81), such cases represent additional violations of Givón's (1983) claims.

[7]I use the cover-term "subject-switching" in order to make a distinction between the phenomenon investigated here and canonical "switch-reference" , where a verb in one clause is overtly marked to indicate whether its subject is identical with, or different to, the subject of the following clause (cf. Jacobsen 1967; Haiman and Munro 1983).

[8]As Paul Schachter (p.c.) has suggested, one might expect to encounter more cases of subject-ellipsis in contexts where the referents controlled differing grammatical gender. I scrutinized one story in the written corpus -- Imam (1970:15-20) -- in which a feminine participant figured prominently, and discovered that for the new subject/same sentence category at least, the skewings are in the predicted direction -- only one NP token (10.0%) vs. a total of nine (90.0%) ellipted subjects, compared with the overall tallies of 41.9% (lexical NP subjects) and 58.1% (ellipted subjects).

[9]Of the 261 tokens of elliptical subject-switches occurring in the written corpus, I encountered only one subject-switch point -- (9e) -- where reference was not immediately apparent. None of my Hausa-speaking friends,

however, had any such momentary difficulties of interpretation.

[10]Robert Kirsner (p.c.) has also suggested that in face-to-face contexts, the fact that the speaker can also provide gestural information may also be relevant.

[11]See Brown and Yule (1983:95-100) for a concise summary of various approaches to the phenomenon of paragraph segmentation.

[12]With regard to the two spoken Pear Film narratives, there were several points at which both speakers hesitated for some considerable time (roughly two seconds or more), before proceeding to a new episode, and I shall provide some examples of the referential choices made at these more obvious breaks in continuity (cf. Chafe 1979b). Because of the general difficulties involved in determining exactly where structural breaks did occur in the spoken texts, however, the following analysis concentrates almost exclusively on the more substantial body of written narrative data.

[13]Chafe (1979b:179) captures the essence of the problem when he writes:

"Our data, then, do not support the hypothesis that a narrative can be unambiguously divided into a fixed number of episodes or paragraphs. They instead suggest that as a speaker moves from focus to focus (or from thought to thought) there are certain points at which there may be a more or less radical change in space, time, character configuration, event structure, or, even, world. Each of these factors may contribute to a processing difficulty at such a point, and each may contribute more or less. The processing difficulty appears in speech as hesitation and is recognized in writing as a paragraph division."

[14]Nichols (1981:10-11) reports that in literary Russian ellipsis -- which conveys no information regarding the identity of the coreferential subject argument -- is not possible across episode-boundaries. Instead, some form of NP or anaphoric pronoun must be used. And Clancy (1980:176n) reports a tendency for ellipsis to be blocked across episode-boundaries in the Japanese Pear Film narratives.

[15]Longacre (1979:118) cites two unrelated languages -- Gurung of Nepal and Sanio-Hiowe of New Guinea -- in which "...we find back reference exclusively within the paragraph and not between paragraphs. Lack of back refe-

rence is indicative, therefore, of a paragraph boundary." Quite how such a powerful, paragraph-level constraint on coreference can work in reality is something of a puzzle to me.

[16]The count includes the subject arguments of performative verbs introducing paragraph-final direct speech quotations (cf. section 5.3).

[17]Russell Schuh (p.c.) has noted, correctly I believe, that because the (15c) clause is still part of a fairly tight sequence of narrative events, one cannot place too much faith in the significance of all the paragraph-breaks Imam (1970) chooses to indicate. This observation does not, however, vitiate the general spirit of my claim that same-subjects tend to be ellipted.

[18]In the interests of perspicuity, I have taken the liberty of paragraph-indenting in those oral narrative contexts where an extended pause seemed to signal the onset of a fresh episode of some kind.

[19]The disparity between the token-totals of Figures 5.2 and 5.5 is due to the following: unlike Figure 5.2, the counts in Figure 5.5 do not include first-mention indefinite NP's; and in addition, the totals in Figure 5.5 include cases of both subject-preservation and subject-switching.

[20]Clancy (1980:172-73), it should be noted, suggests that the shift from nonovert to overt reference at episode-boundaries could be the outcome of two determinants. On the one hand, such a coding escalation might be "listener-oriented", i.e. the speaker manipulates a more explicit device in order to indicate structural boundaries to the listener. On the other hand, it could be that the speaker, wrestling with the cognitive task of retrieving and verbalizing the essence of a new episode, feels the need to reactivate material which might have drifted away from his/her focal consciousness. Cf. too Chafe (1979b) on this matter.

[21]Li and Thompson (1979) report a tendency in Mandarin Chinese texts to use an escalated referential strategy -- usually pronouns in preference to zero-anaphora -- following time-adverbials.

[22]The single occurrence of the demonstrative was used to modify the second lexical mention of the kin-term in question, and so is probably best analyzed as a case of new referent-anchoring (cf. Chapter 4).

[23]Use of the can-marked demonstrative here seems to substantiate the claim advanced in Chapter 4 that this set is generally manipulated "to denote referents in contexts which are highly discontinuous", i.e. mapping entities which are more remote-distal in either a physical-locative or discourse-referential sense.

[24]Comparable examples are reported for Godie, a Kru language spoken in the Ivory Coast (Marchese 1982:5-6); and Clancy (1980:175) notes cases in English.

[25]Since "turn-taking" conversational exchanges (cf. Sacks et al. 1974) are usually paragraph-indented, we are, in a rather literal sense, still working with p-marked breaks.

[26]In terms of cognitive processing, the effect of intervening direct speech is perhaps comparable to the impact on referential choice of such phenomena as "digressions" and "world-shifts" as discussed in Chafe (1979b) and Clancy (1980). As far as other languages are concerned, Nichols (1981:11-12) reports that, in literary Russian, nonovert marking of a post-quotation boundary same-subject ("theme"), i.e. using "true" ellipsis, is impermissible -- some form of pronoun or NP must be used. And Duncan's (1982:15-16) data on Chamorro point to a correlation between direct speech closure and the following use of full NP's to encode interactant switches.

CHAPTER 6

CONCLUDING REMARKS

The primary thrust of this dissertation has, I believe, been largely consistent with the traditional concern of the "dirty-hands" descriptive linguist -- to attempt to account for the manner in which linguistic elements are exploited in communication, i.e. how encoders package the referential information they transmit. In verbalizing the characters and objects which make up a narrative structure, an encoder is obliged to make decisions regarding the referential forms which s/he considers most appropriate, and I have tried to determine some, at least, of the more influential factors conditioning such decisions. Discourse analysis is essentially concerned with describing, via statistically-based generalizations, observable, cross-text regularities, and with establishing the frequency with which a given linguistic item is encountered in a specifiable context. We have noted, on more than one occasion that this kind of discourse-based investigation cannot yield the types of absolute, invariant "rules" which are characteristic of sentence-based analyses; rather, there are discernible tendencies with greater and smaller numbers of exceptions. Several issues remain

unresolved, particularly, I believe, the problem of the referential competition between the two deictic markers discussed in Chapters 4 and 5. I hope, however, that these losses are outweighed by some of the gains and insights achieved in other areas.

Finally, I believe that some of the documented claims about the relationship between discourse-salience and morphological coding have some interesting implications for Haiman's (1980, 1983) proposals concerning the "diagrammatic iconicity" of grammars. The essence of Haiman's position is that the exploitation of linguistic forms is less arbitrary than has hitherto been assumed, and that there is a discernible iconic link between the physical shape of linguistic elements and the nature of the messages these structures signal. Haiman's hypothesis has, moreover, been profitably used by some linguists as a basis for investigating and explaining various iconic structures in languages (cf. Hopper and Thompson 1983), and I believe we have noted two significant phenomena, in this dissertation, which may also be reasonably analyzed as instances of iconically-motivated choices -- the strong tendency to use an Indefinite Specifier and a full direct object pronoun, in preference to zero-marking, for highly-salient human discourse-participants. At the same time, it is important to note that the facts of the pronoun:zero anaphor distribution in the direct object position are

direct violations of Haiman's additional claim that "reduction of form is an ECONOMICALLY motivated index of familiarity, not an iconically motivated index" (1983:802), this being simply a restatement of Givón's (1983) hypothesis that the formal devices used to mark discourse "topics" co-vary, in their phonological bulk and complexity, with the (dis)continuity/(un)predictability of the referent. The differential coding chosen for equally predictable, equally familiar referents, with humans assigned a much greater proportion of full pronouns in objectively comparable linguistic environments, must surely represent an index of the greater conceptual complexity and pragmatic-disccurse salience of such arguments.

APPENDIX I

Data for Figure 2.1: Written Narrative Counts and Percentages of IS-marked Indefinites with respect to Semantic Class and Syntactic Status.

Semantic class	Syntactic category			
	Subject		Nonsubject	
	No.	%	No.	%
Human [+IS]	106/122	86.9	76/131	58.0
Human [-IS]	16/122	13.1	55/131	42.0
Animal [+IS]	19/33	57.6	7/31	22.6
Animal [-IS]	14/33	42.4	24/31	77.4
Inanimate [+IS]	3/7	42.9	101/432	23.4
Inanimate [-IS]	4/7	57.1	331/432	76.6

APPENDIX II

Data for Figure 2.2: Average Number of Subsequent Mentions of Indefinites with respect to IS-marking, Semantic Class, and Syntactic Status (Written Narratives).

Semantic class	Syntactic category			
	Subject		Nonsubject	
	Mean # of mentions		Mean # of mentions	
Human [+IS]	19.0	(2010/106)	8.7	(661/76)
Human [-IS]	2.0	(32/16)	6.2	(339/55)
Animal [+IS]	15.8	(301/19)	8.9	(62/7)
Animal [-IS]	3.9	(55/14)	4.1	(99/24)
Inanimate [+IS]	0.7	(2/3)	1.7	(174/101)
Inanimate [-IS]	0.8	(3/4)	0.7	(228/331)

APPENDIX III

Data for Figure 3.1: Written Narrative Distribution of the Three Referential Categories with respect to the Number of Clauses Intervening between Successive Mentions of a Referent.

No. of Clauses	Referential Type					
	Zero		Pronoun		NP	
0			22 / 221	9.9%		
1	515 / 636	81.0%	113 / 221	51.1%	28 / 315	8.8%
2-4	98 / 636	15.3%	61 / 221	27.6%	84 / 315	26.7%
5-10	12 / 636	1.9%	17 / 221	7.7%	102 / 315	32.4%
11-20	10 / 636	1.6%	6 / 221	2.7%	59 / 315	18.7%
21-30	1 / 636	0.2%	1 / 221	0.5%	15 / 315	4.8%
31+			1 / 221	0.5%	27 / 315	8.6%

APPENDIX IV

Data for Figure 3.2: Spoken Narrative Distribution of the Three Referential Categories with respect to the Number of Clauses Intervening between Successive Mentions of a Referent.

No. of Clauses	Referential Type					
	Zero		Pronoun		NP	
0			13 / 88	14.8%		
1	153 / 200	76.5%	47 / 88	53.4%	21 / 85	24.7%
2-4	32 / 200	16.0%	27 / 88	30.7%	23 / 85	27.1%
5-10	10 / 200	5.0%	1 / 88	1.1%	25 / 85	29.3%
11-20	4 / 200	2.0%			10 / 85	11.8%
21-30					2 / 85	2.4%
31+					4 / 85	4.7%

APPENDIX V

Data for Figure 3.3: Written Narrative Distribution of the Three Referential Categories with respect to the Number of Confusers Intervening between Successive Mentions of a Referent.

No. of Confusers	Referential Type					
	Zero		Pronoun		NP	
0	591 / 636	92.9%	196 / 221	88.7%	170 / 315	53.9%
1	44 / 636	6.9%	24 / 221	10.8%	89 / 315	28.3%
2-4	1 / 636	0.2%	1 / 221	0.5%	55 / 315	17.5%
5+					1 / 315	0.3%

APPENDIX VI

Data for Figure 3.4: Spoken Narrative Distribution of the Three Referential Categories with respect to the Number of Confusers Intervening between Successive Mentions of a Referent.

No. of Confusers	Referential Type					
	Zero		Pronoun		NP	
0	189 / 200	94.5%	80 / 88	90.9%	54 / 85	63.6%
1	11 / 200	5.5%	8 / 88	9.1%	24 / 85	28.3%
2-4					6 / 85	7.1%
5+					1 / 85	0.3%

APPENDIX VII

Data for Figure 3.5: Written Narrative Distribution of Direct Object Pronoun Anaphora and Direct Object Zero Anaphora with respect to Semantic Class

Class	Anaphoric device	
	DO Zero	DO Pronoun
HUMAN	15 / 181 = 8.3%	166 / 181 = 91.7%
ANIMAL	26 / 43 = 60.5%	17 / 43 = 39.5%
INANIMATE	168 / 192 = 87.5%	24 / 192 = 12.5%

APPENDIX VIII

Data for Figure 3.6: Distribution of Direct Object Pronouns and Direct Object Zero Anaphors in Written Narratives: Numbers and Percentages according to Semantic Class.

Class	Anaphoric device	
	DO Zero	DO Pronoun
HUMAN	15 / 209 = 7.2%	166 / 207 = 80.2%
ANIMAL	26 / 209 = 12.4%	17 / 207 = 8.2%
INANIMATE	168 / 209 = 80.4%	24 / 207 = 11.6%

APPENDIX IX

Data for Figures 5.1 and 5.2: Percentages of Referential Forms Chosen for Preserving and Switching Subject Reference within Sentences and across Sentence Boundaries

	Same Subject				New Subject			
	Same sentence		New sentence		Same sentence		New sentence	
Written	%	No.	%	No.	%	No.	%	No.
Noun phrase			3.6	5	41.9	121	71.1	229
Ellipsis	100.0	602	96.4	132	58.1	168	28.9	93
		602 (81.5%)		137 (18.5%)		289 (47.3%)		322 (52.7%)
Oral								
Noun phrase			6.5	3	44.1	19	58.5	24
Ellipsis	100.0	106	93.5	43	55.9	24	41.5	17
		106 (69.7%)		46 (30.3%)		43 (51.2%)		41 (48.8%)

APPENDIX X

Data for Figure 5.3: Written Narrative Percentages of Differential NP Configurations Chosen for Switching Subject Reference within and across Sentence Boundaries.

Category	Same Sentence		New Sentence	
N + demonstrative	11 / 42	26.2%	31 / 42	73.8%
N + determiner	9 / 26	34.6%	17 / 26	65.4%
Simple N	66 / 167	39.5%	101 / 167	60.5%
Proper name	25 / 94	26.6%	69 / 94	73.4%
N + possessor	10 / 21	47.6%	11 / 21	52.4%

APPENDIX XI

Data for Figure 5.4: Written Narrative Percentages of Nominal and Elliptical Subjects Chosen where Prior Mention is within or across Paragraph-Marked Boundaries.

Category	Same Paragraph as Prior Control		Different Paragraph from Prior Control	
Noun Phrase	88 / 281	31.3%	193 / 281	69.7%
Ellipsis	958 / 994	96.4%	36 / 994	3.6%

APPENDIX XII

Data for Figure 5.5: Written Narrative Percentages of Differential NP Configurations Chosen where Prior Mention is within or across Paragraph-Marked Boundaries.

Category	Same Paragraph as Prior Control		Different Paragraph from Prior Control	
N + demonstrative	12/88	13.6%	30/193	15.5%
N + determiner	5/88	5.7%	23/193	11.9%
Simple noun	44/88	50.0%	56/193	29.0%
Proper name	26/88	29.6%	70/193	36.3%
N + possessor	1/88	1.1%	14/193	7.3%

APPENDIX XIII

Data for Figure 5.6: Written Narrative Percentages of Differential NP Configurations Chosen for Immediate Post-Direct Speech Subjects.

Category	Following "Direct Speech" Quotation Boundary		Percentage of Total Category Tokens	
N + DEM	4/86	4.7%	4/42	9.5%
N + DET	13/86	15.1%	13/28	46.4%
Bare N	40/86	46.5%	40/100	40.0%
Proper name	25/86	29.0%	25/96	26.0%
N + POSS	4/86	4.7%	4/15	26.7%

Bibliography

Abraham, Roy C. 1959. The Language of the Hausa People. London: University of London Press.

Abraham, Roy C. 1962. Dictionary of the Hausa Language, 2nd ed. London: University of London Press.

Bentivoglio, Paola. 1983. Topic continuity and discontinuity in discourse: A study of spoken Latin-American Spanish. In (ed.) Talmy Givón 1983, pp. 255-311.

Bernardo, Robert. 1980. Subjecthood and consciousness. In (ed.) Wallace L. Chafe 1980b, pp. 275-99.

Brown, Gillian, and George Yule. 1983. Discourse Analysis (Cambridge textbooks in Linguistics). Cambridge: Cambridge University Press.

Chafe, Wallace L. 1974. Language and consciousness. Language 50:111-53.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, and point of view. In (ed.) Charles N. Li 1976, pp. 25-55.

Chafe, Wallace L. 1979a. Integration and involvement in speaking, writing, and oral literature. To appear in Tannen 1981.

Chafe, Wallace L. 1979b. The flow of thought and the flow of language. In (ed.) Talmy Givón 1979a, pp. 159-81.

Chafe, Wallace L. 1980a. The deployment of consciousness in the production of a narrative. In (ed.) Wallace L. Chafe 1980b, pp. 9-50.

Chafe, Wallace L. 1980b (ed.). The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production. Norwood, N.J.: Ablex.

Clancy, Patricia M. 1980. Referential choice in English and Japanese narrative discourse. In (ed.) Wallace L. Chafe 1980b, pp. 127-202.

Cowan, J Ronayne, and Russell G. Schuh. 1976. Spoken Hausa. Ithaca: Spoken Language Services.

Du Bois, John W. 1980a. Introduction -- The search for a cultural niche: showing the pear film in a Mayan Community. Preface to (ed.) Wallace L. Chafe 1980b.

Du Bois, John W. 1980b. Beyond definiteness: The trace of identity in discourse. In (ed.) Wallace L. Chafe 1980b, pp. 203-74.

Du Bois, John W. 1981. The Sacapultec language. Berkeley: Unpublished Ph.D. dissertation.

Duncan, Janine S. 1982. Referential strategies and preferred clause structure in Chamorro. UCLA master's thesis.

Duranti, Alessandro, and Elinor Ochs. 1979. Left-dislocation in Italian conversation. In (ed.) Talmy Givón 1979a, pp. 377-416.

Fillmore, Charles J. 1977. Topics in lexical semantics. In (ed.) Roger W. Cole 1977, Current Issues in Linguistic Theory, pp. 76-138. Bloomington, Indiana: Indiana University Press.

Fries, Peter H. 1980. On the status of theme in English: Arguments from discourse. To appear in Forum Linguisticum.

García, Erica C. 1975. The Role of Theory in Linguistic Analysis: The Spanish Pronoun System. Amsterdam: North Holland.

García, Erica C. 1983. Context dependence of language and of linguistic analysis. In (ed.) Flora Klein-Andreu 1983a, pp. 181-207.

Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In (ed.) Charles N. Li 1976, pp. 149-88.

Givón, Talmy. 1977. Definiteness and referentiality. In Joseph Greenberg, C. Ferguson, and E. Moravcsik (eds.), Universals of Human Language, pp. 291-330. Stanford: Stanford University Press.

Givón, Talmy. 1979a (ed.). Discourse and Syntax. New York: Academic Press.

Givón, Talmy. 1979b. On Understanding Grammar. New York: Academic Press.

Givón, Talmy. 1980. Ute Reference Grammar. Ignacio, Colorado: Ute Press.

Givón, Talmy. 1981. Topic continuity: The functional domain of switch-reference. To appear in (eds.) Haiman and Munro 1983.

Givón, Talmy. 1982. Logic vs. pragmatics, with human language as the referee: toward an empirically viable epistemology. Journal of Pragmatics 6.2:81-133.

Givón, Talmy. 1983 (ed.). Topic Continuity in Discourse: A Quantitative Cross-Language Study. (Typological Studies in Language, vol. 3). Amsterdam: John Benjamin.

Greenberg, Joseph H. 1977. Niger-Congo noun class markers: prefixes, suffixes, both or neither. In (eds.) Martin Mould and Thomas J. Hinnebusch, Papers from the Eighth Conference on African Linguistics, pp. 97-104. Studies in African Linguistics, Suppl. 7.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In (ed.) Joseph H. Greenberg, Universals of Human Language, vol. 2, pp. 47-82. Stanford, California: Stanford University Press.

Grice, H. Paul. 1975. Logic and conversation. In (eds.) P. Cole and J. Morgan, Syntax and Semantics 3: Speech Acts, pp. 41-58. New York: Academic Press.

Grimes, Joseph E. 1975. The thread of discourse. Janua Linguarum 207. The Hague: Mouton.

Grimes, Joseph E. 1978. Papers on Discourse. Dallas: Summer Institute of Linguistics, Inc.

Haiman, John. 1980. The iconicity of grammar: Isomorphism and motivation. Language 56(3):515-40.

Haiman, John. 1983. Iconic and economic motivation. Language 59(4):781-819.

Haiman, John, and Pamela Munro. 1983 (eds.). Switch-Reference and Universal Grammar. (Typological Studies in Language, vol. 2). To appear. Amsterdam: John Benjamins.

Hetzron, Robert. 1971. Presentative function and presentative movement. Studies in African Linguistics, Suppl. 2, pp. 79-105.

Hinds, John. 1977. Paragraph structure and pronominalization. Papers in Linguistics 10:77-99.

Hinds, John. 1978a. Levels of structure within the paragraph. Berkeley Linguistic Society 4:598-609.

Hinds, John. 1978b (ed.). Anaphora in Discourse. Alberta: Linguistic Research, Inc.

Hinds, John. 1979. The paragraph as a grammatical unit. In (ed.) Talmy Givón 1979a, pp. 115-34.

Hinds, John. 1983. Topic continuity in Japanese. In (ed.) Talmy Givón 1983, pp. 43-93.

Hopper, Paul J. 1979. Aspect and foregrounding in discourse. In (ed.) Talmy Givón 1979a, pp. 213-41.

Hopper, Paul J., and Sandra A. Thompson. 1983. The discourse basis for lexical categories in universal grammar. To appear in Language 60.4.

Jacobsen, William H., Jr. 1967. Switch-reference in Hokan-Coahuiltecan. Studies in Southwestern Ethnolinguistics, ed. by Dell Hymes, pp. 238-63. The Hague: Mouton.

Jaggar, Philip J. 1978. And what about...? -- Topicalisation in Hausa. Studies in African Linguistics 9:69-81.

Jaggar, Philip J. 1981. The Kano dialect of Hausa in the Polyglotta Africana. Afrika und Übersee 64:47-58.

Jaggar, Philip J. 1982a. Varieties of passive in Hausa. Papers from the Twelfth Annual Conference on African Linguistics, ed. by William R. Leben. Studies in African Linguistics, Suppl. 8:73-77.

Jaggar, Philip J. 1982b. The two perfective aspects of Hausa and their roles in the flow-control of narrative structures. MS, Department of Linguistics, UCLA.

Jaggar, Philip J. 1983. Some dimensions of topic-NP continuity in Hausa narrative. In (ed.) Talmy Givón 1983, pp. 365-424.

Keenan, Edward L. 1976. Towards a universal definition of 'subject'. In (ed.) Charles N. Li 1976, pp. 247-301.

Kirsner, Robert S. 1979. Deixis in discourse: An exploratory quantitative study of the Modern Dutch demonstrative adjectives. In (ed.) Talmy Givón 1979a, pp. 355-75.

Kirsner, Robert S., and Vincent J. van Heuven. 1980. On the opposition between deze (dit) and die (dat) in written Dutch: a discriminant analysis. In (eds.) Saskia Daalder and Marinel Gerritsen, Linguistics in the Netherlands, pp. 101-108. Amsterdam: North-Holland Publishing Co.

Klatzky, R. 1975. Human Memory: Structures and Processes. San Francisco, California: W. H. Freeman.

Klein-Andreu, Flora. 1983a (ed.). Discourse Perspectives on Syntax. New York: Academic Press.

Klein-Andreu, Flora. 1983b. Grammar in style: Spanish adjective placement. In (ed.) Flora Klein-Andreu 1983a, pp. 143-79.

Kraft, Charles H., and A.H.M. Kirk-Greene. 1973. Hausa. (Teach Yourself Books). London: The English Universities Press.

Li, Charles N. 1976 (ed.). Subject and Topic. New York: Academic Press.

Li, Charles N., and Sandra A. Thompson. 1976. Subject and topic: a new typology of language. In (ed.) Charles N. Li 1976, pp. 457-90.

Li, Charles N., and Sandra A. Thompson. 1979. Pronouns and zero-anaphora in Chinese discourse. In (ed.) Talmy Givón 1979a, pp. 311-35.

Longacre, Robert E. 1979. The paragraph as a grammatical unit. In (ed.) Talmy Givón 1979a, pp. 115-34.

Marchese, Lynell. 1982. Pronominalization and the appearance of full noun referents in Godie discourse. To appear in (ed.) G. Dimmendaal, Current Approaches to African Linguistics, vol. 3.

McConvell, Patrick. 1973. Cleft sentences in Hausa? A syntactic study of focus. London: unpublished Ph.D. dissertation, University of London.

Newman, Paul. 1972. Study Kanakuru, understand Hausa. Harsunan Nijeriya 2:1-13. Kano, Nigeria: Centre for the Study of Nigerian Languages.

Newman, Paul, and Russell G. Schuh. 1974. The Hausa aspect system. Afroasiatic Linguistics 1(1):1-39.

Nichols, Johanna. 1981. The grammatical marking of theme in literary Russian. To appear in (eds.) Richard D. Brecht and Michael Flier, Issues in Russian Morphosyntax.

Norman, D. A. 1969. Memory and Attention. New York: Wiley.

Parsons, F.W. 1960. The verbal system in Hausa. Afrika und Übersee 44:1-36.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. Language 50:696-735.

Schachter, Paul. 1966. A generative account of Hausa NE/CE. Journal of African Languages 5:34-54.

Schachter, Paul. 1973. Focus and relativization. Language 49:19-46.

Schuh, Russell G. 1977. Bade/Ngizim determiner system. Afroasiatic Linguistics 4(3):1-74.

Schuh, Russell G. 1983. The evolution of determiners in Chadic. In (eds.) Ekkehard Wolff and Hilke Meyer-Bahlburg, Studies in Chadic and Afroasiatic Linguistics, pp. 157-210. Hamburg: Helmut Buske Press.

Skinner, Neil A. 1974. Hausa wani/wata/wa'dansu and its semantic functions. Third Annual Conference on African Linguistics, ed. by Erhard Voeltz. Indiana University Publications, African Series, vol. 7, pp. 251-57.

Tannen, Deborah. 1980. A comparative analysis of oral narrative strategies: Athenian Greek and American English. In (ed.) Wallace L. Chafe 1980b, pp. 51-87.

Tannen, Deborah. 1981 (ed.). Spoken and Written Language: Exploring Orality and Literacy. Norwood, NJ: Ablex. To appear.

Tannen, Deborah. 1982. Oral and literate strategies in spoken and written narratives. Language 58(1):1-21.

Wald, Benji. 1979. The development of the Swahili object marker: A study of the interaction of syntax and discourse. In (ed.) Talmy Givón 1979a, pp. 505-24.

Wald, Benji. 1983. Referents and topic within and across discourse units: Observations from current vernacular English. In (ed.) Flora Klein-Andreu 1983a, pp. 91-116.

Zubin, David A. 1979. Discourse function of morphology: The focus system in German. In (ed.) Talmy Givón 1979a, pp. 469-504.

Published Data Sources

Imam, Abubakar. 1966 [1939]. Magana Jari Ce, vol. 3. Zaria, Nigeria: Northern Nigerian Publishing Company.

Imam, Abubakar. 1970 [1939]. Magana Jari Ce, vol. 2. Zaria, Nigeria: Northern Nigerian Publishing Company.

Imam, Abubakar. 1971 [1937]. Magana Jari Ce, vol. 1. Zaria, Nigeria: Northern Nigerian Publishing Company.