

CONTINUITY AND GIVEN-NEW STATUS OF DISCOURSE REFERENTS
IN ADZERA ORAL NARRATIVE

by

DAVID EDWARD HOWARD

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF ARTS IN LINGUISTICS

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2002

Copyright © by David Edward Howard 2002
All Rights Reserved

ACKNOWLEDGMENTS

There are many people to whom I am grateful for their help and encouragement in this project. First, I want to thank my supervising professor, Dr. Laurel Stvan, for the many hours she spent reading over drafts and providing helpful feedback, insights, direction, and encouragement. Thanks are also due to the other members of my committee, Dr. Shin Ja Hwang and Dr. Donald Burquest, for giving helpful insights and guidance in various areas and asking important questions. Thank you to the G.I.A.L. library and staff for providing me with numerous linguistic resources.

I want to thank the government of Papua New Guinea (PNG), for allowing my family and me to live and work in the country. I am grateful to the PNG Branch of SIL—under whose auspices we work—who provided scholarship funds to help with tuition and fees for my final two semesters at UTA. In particular, I thank Bob Bugenhagen for giving guidance on the thesis topic, and Morris and Wendy Johnson, who have helped and supported us in many ways, including providing Adzera language resources. I am grateful also to our financial and prayer partners, who enable us by their generosity to live and work in PNG.

I am particularly indebted to my Adzera friends who have helped me to learn their language. There are too many to list here, but some deserve special mention. I thank Gadan Ayam and Wari Muruk for providing the texts on which this study is based. Wari also helped me to transcribe and gloss many texts. Numerous others have provided me with texts, which I hope to use in future studies. Special thanks go to my partners in translation:

Tengeng Bibuae, Sakingwaga Rubus, Painsa Piriang, and Ngawi Geasa, for helping me to understand their language on a deeper level and for providing insights on particular issues. Thanks also to Tengeng, his father Bibuae, and their family for allowing us to live on their land as we undertake our work. *Dajki tsira' biran!*

I especially want to thank my wife, Susan, who has not only been a constant source of encouragement, but who has also worked so hard to keep our household running in order to free me to concentrate on my studies. She deserves a great deal of credit. I also want to thank my children—Katie, Melissa, and Samuel—who put up with me when I locked myself away from the family and could not give them as much time and attention as all of us would have liked. Finally, I thank my Lord for calling me and equipping me to work as a linguist and translator in PNG, and for giving me the strength for this project and everything I do.

November 18, 2002

ABSTRACT

CONTINUITY AND GIVEN-NEW STATUS OF DISCOURSE REFERENTS

IN ADZERA ORAL NARRATIVE

Publication No. _____

David Edward Howard, M.A.

The University of Texas at Arlington, 2002

Supervising Professor: Laurel Smith Stvan

This thesis analyzes the encoding of discourse referents in oral narratives in Adzera, an SVO Austronesian language of Papua New Guinea. First, it classifies overt referents on a given-new continuum according to Prince's (1981) taxonomy of assumed familiarity, looking at the forms and positions used to introduce new entities into a discourse. Then it applies Givón's (1983, 1995) quantitative approach to measure the continuity of referents, whether overtly manifested or semantically implied as verb arguments.

This study largely confirms the findings and predictions of Prince and Givón regarding the introduction of new information and the relative continuity of various forms of reference and grammatical positions. Zero anaphora encode the majority of highly accessible subjects that would be encoded by pronouns in English, even though Adzera has no subject-verb agreement. Of particular interest are instances of first mention by pronominal forms, which appear to function in place of a passive construction, which Adzera lacks.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
ABSTRACT	vi
LIST OF FIGURES	xii
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xv
Chapter	
1. INTRODUCTION	1
1.1 Purpose and Scope of this Study	1
1.2 The Adzera Language	2
1.3 Previous Research on Adzera Language and Culture	3
1.4 Adzera Culture and the Place of Oral Narrative	4
1.5 Data for the Current Study	6
1.6 Expected Findings	8
1.7 Organization of the Thesis	9
2. AN OVERVIEW OF ADZERA GRAMMAR	10
2.1 Introduction	10
2.2 Phonemes and Graphemes	11
2.3 Word Order	13

2.4 Noun Phrases	16
2.4.1 NP Constituents	16
2.4.2 Definite and Indefinite NPs	17
2.4.3 Possession	19
2.4.4 Pronouns	22
2.4.5 Relative Clauses	23
2.5 Verb Phrases	25
2.5.1 Verbal Morphology	26
2.5.2 Ellipsis of Subject and Object	27
2.5.3 Serial Verbs	29
2.5.4 Clausal Complements	30
3. THE THEORETICAL FRAMEWORK FOR THE PRESENT STUDY	36
3.1 Introduction	36
3.2 Grice's Maxims	37
3.3 Given vs. New Information	38
3.3.1 Given and New Defined	38
3.3.2 Propositional vs. Lexical or Referential Information	41
3.3.3 Identifiability	43
3.4 Prince's Taxonomy of Assumed Familiarity	44
3.4.1 Integrating the Given-New Dichotomies	44
3.4.2 Evoked Entities	45
3.4.3 New Entities	45

3.4.4 Old and New Entities in Terms of Hearer-status and Discourse-status	46
3.4.5 Inferable Entities	47
3.5 Topic Continuity	52
3.6 Participant Reference	55
3.7 Related Issues	56
4. ASSUMED FAMILIARITY OF DISCOURSE REFERENTS	57
4.1 Introduction	57
4.2 Potential Problems of Assumed Familiarity	57
4.3 Apparent Problems of Assumed Familiarity in an Adzera Text	60
4.4 A Classification of Referents in the Adzera Texts	68
4.4.1 Methodology	68
4.4.2 Specific Examples of Referent Classes	72
4.4.2.1 Examples of Brand-new Referents	72
4.4.2.2 Examples of Brand-new Anchored Referents	73
4.4.2.3 Examples of Unused Referents	76
4.4.2.4 Examples of Inferable Referents	77
4.4.2.5 Examples of Containing Inferable Referents	78
4.4.2.6 Examples of Evoked Referents	79
4.4.2.7 Examples of Situationally Evoked Referents	80
4.4.3 Referents Within Quotations	82
4.4.4 Quantitative Distribution of Referent Classes	83

4.5 Summary	90
5. A QUANTITATIVE ANALYSIS OF TOPIC/PARTICIPANT CONTINUITY	92
5.1 Introduction	92
5.2 Methodology	93
5.2.1 Measurements	93
5.2.2 Counting Clauses	96
5.2.3 Defining Topics/Participants and Counting Their Occurrences ..	97
5.2.4 Counting and Computing Methods	102
5.2.5 Illustration of the Methodology on a Portion of Text	103
5.3 Predictions	107
5.4 Results	111
5.4.1 Overview	111
5.4.2 Accessibility and Form	111
5.4.2.1 Accessibility of Animate Participants	111
5.4.2.2 Accessibility of Inanimate Topics	117
5.4.3 Topic Persistence and Form	120
5.4.3.1 Persistence of Animate Participants	120
5.4.3.2 Persistence of Inanimate Topics	124
5.4.3.3 Persistence and the Influence of Animacy	125
5.4.4 Continuity and Grammatical Position	125
5.4.4.1 Accessibility and Grammatical Position	125
5.4.4.2 Topic Persistence and Grammatical Position	129

5.5 Summary	131
6. FINDINGS AND CONCLUSIONS	133
6.1 Summary of Findings	133
6.1.1 Assumed Familiarity of Overt Discourse Referents and Their Distribution	133
6.1.2 Continuity of Discourse Referents	135
6.2 Conclusions and Applications	137
6.3 Future Research Directions	139
 Appendix	
A. TEXT A: OPPOSITION TO MISSION WORK	142
B. TEXT B: MARAMAI SPEARS A CROCODILE	155
C. RELATIVE CLAUSES, COMPOUNDS, AND COMPLEMENT CLAUSES	164
REFERENCES	168
BIOGRAPHICAL INFORMATION	175

LIST OF FIGURES

Figure	Page
1. The family tree of the Adzera language	2
2. Prince's taxonomy of assumed familiarity	50
3. Prince's familiarity scale	51
4. Givón's scale of phonological size	107

LIST OF TABLES

Table	Page
1. Adzera Consonants	12
2. Demonstratives	18
3. Personal Pronouns	22
4. Preverbal Morphemes and Prefixes	26
5. Prince's Categories in Terms of Hearer- and Discourse-status	47
6. General Distribution of Referents	84
7. Total Referents in Text A	87
8. Total Referents in Text B	87
9. Animate Referents in Text A	88
10. Animate Referents in Text B	88
11. Inanimate Referents in Text A	89
12. Inanimate Referents in Text B	89
13. Interpreting the Measurements	95
14. Contrasting the Two Approaches	101
15. Summary of Sample References and Their Continuity Scores	107
16. Possessor Marking	109
17. Average RD and PI Values for Text A (Animate Referents)	112

18. Adjusted Average RD and PI Values for Text A (Animate Referents)	112
19. Percentages of High and Low RD and PI Values for Text A (Animate Referents)	113
20. Percentages of High and Low RD and PI Values for Text B (Animate Referents)	114
21. Percentages for RD and PI of Possessors in Both Texts (Animate Referents)	115
22. RD and PI by Text and Reference Form in Three Tiers (Animate Referents)	116
23. Accessibility (RD and PI) for Both Texts (Animate Referents)	117
24. Accessibility (RD and PI) of Inanimate Topics	118
25. A Comparison of Accessibility of Inanimate Topics in Both Texts	119
26. Persistence of Animate Participants by Text and Form of Reference	121
27. Persistence of Main vs. Supporting Characters	122
28. Persistence of Inanimate Topics by Text and Form of Reference	124
29. Referential Distance of Animate Subjects vs. Non-subjects	126
30. Referential Distance of Subjects vs. Non-subjects (Animate and Inanimate Combined)	128
31. RD Values of Subject vs. Non-subject by Referent Type	128
32. Topic Persistence of Animate Subjects vs. Non-subjects	130
33. Topic Persistence of Subjects vs. Non-subjects (Animate and Inanimate Combined)	131

LIST OF ABBREVIATIONS

CMPLT	completive aspect	R	realis mood
Comp	complement marker	RECIP	reciprocal
CONT	continuous aspect	RC	relative clause
Dem	demonstrative	S	subject (marks word order)
DO	direct object	sg	singular
EXCL	exclamation	SPEC	specifier
FUT	future tense	SOV	subject-object-verb
HORT	hortative mood	SVO	subject-verb-object
IMP	imperative mood	V	verb
INTENS	intensifier	VP	verb phrase
IR	irrealis mood	1pl.ex	first person plural exclusive
LOC	locative	1pl.in	first person plural inclusive
N	noun	1pss	first person possessor
NEG	negative	1sg	first person singular
NMZ	nominalizer	2pl	second person plural
NP	noun phrase	2pss	second person possessor
O	object (marks word order)	2sg	second person singular
Obj	object marker	3obj	third person object
Obl	oblique	3pl	third person plural
OSV	object-subject-verb	3pss	third person possessor
pl	plural	3sg	third person singular
PREP	preposition	∅	zero anaphor (ellipsis)
prev	previous mention/presence	(um)/(uh)	hesitation (no meaning)
PRT	particle		
PURP	purpose		

CHAPTER 1

INTRODUCTION

1.1 Purpose and Scope of this Study

The purpose of this study is to examine the forms of reference that encode discourse referents in Adzera oral narrative, and to measure the continuity of those referents. In particular, it shows how narrators refer to given vs. new entities as well as topics of high vs. low continuity. It builds upon the discourse studies done in information structure theory, particularly Prince 1981, while employing the quantitative approach to topic continuity pioneered by Givón (1983, 1995).

This type of approach has not been applied previously to the Adzera language, so this study will make a significant contribution to our understanding of Adzera discourse, which, as a whole, also remains unexplored. It will demonstrate what forms of reference are most appropriate in different textual environments. Combined with similar studies of other genres, it will be particularly useful for faithfully translating foreign texts into Adzera in a coherent, cohesive, and natural way, and helping Adzera authors to write as naturally as they speak. It will also be useful for comparing the reference strategies of different languages.

Although the focus of this study is not syntax, the paper includes an overview of certain aspects of Adzera grammar in order to provide helpful background information for the reader. The remainder of the paper focuses on the discourse level, first presenting the

information structure framework, and then applying elements of that framework—in particular the approaches of Prince and Givón—to transcribed texts of Adzera oral narrative.

1.2 The Adzera Language

Adzera is the mother tongue of approximately 30,000¹ people, who live in the upper Markham Valley of Morobe Province in Papua New Guinea. The language belongs to the Austronesian family, and its “family tree” is represented in figure 1:

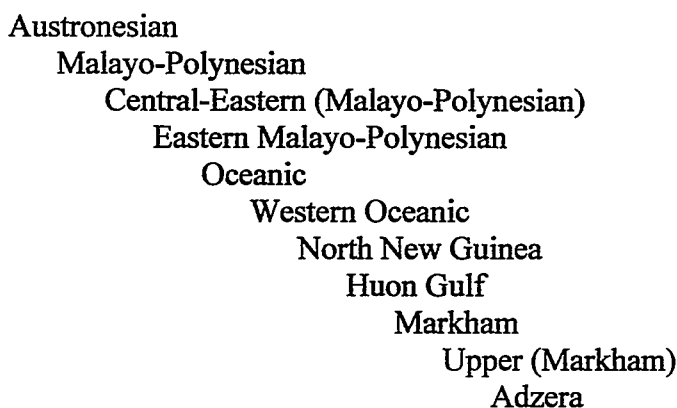


Figure 1. The family tree of the Adzera language. (Data from B. Grimes 2000.)

Its speech area is bordered geographically by several related languages in the Markham subfamily, including Wampar, Wampur, Mari, Sukurum, Sarasira, and North Watut (S. Holzknrecht 1989), as well as some unrelated “Papuan” languages found in the mountainous regions.

¹ This estimate is based on community counts from the Papua New Guinea 2000 National Census, available in electronic form from the National Statistical Office, Port Moresby. Using the provided Community Profile System computer program written by Harry Lode, July 2002, I performed a population count for the Adzera villages. Village information from K. Holzknrecht 1978, S. Holzknrecht 1989, and Adzera speakers, was a great help in identifying the Adzera villages for the program to count. Even so, I may have excluded a few Adzera villages with which I was unfamiliar or included a few small non-Adzera villages. The program calculated a population of 29,687, based on my village selections. This population figure does not include Adzera speakers living in other areas, such as in settlements in Lae or the lower Markham Valley.

Adzera is a head-first language with rigid SVO word order. Its morphological typology is primarily isolating. Verbs are not marked for agreement or case. Although there is some verbal morphology to indicate mood, tense, and aspect, many of the finer distinctions are often determined by context. There is only person-agreement morphology on possessed nouns, and most pronouns do not vary with respect to grammatical position. I will cover more specifics of the language in the next chapter.

1.3 Previous Research on Adzera Language and Culture

According to K. Holzknacht (1978:i) the first person to attempt to describe the Adzera language was the Rev. Fritz Oertel, a Lutheran missionary from Germany. He made visits into the Adzera area before 1917, and was successful in getting Adzera language informants to come to the mission station in the lower Markham Valley. In 1918 he moved to a new permanent mission station within Adzera territory. He wrote an initial alphabet and collected language material, which he shared with Dr. Otto Dempwolff, who wrote a tentative analysis of Adzera (Dempwolff c.1928) based on Oertel's material. After Oertel's death in 1938, all his linguistic material on Adzera was destroyed in World War II, except for a small amount of translated church literature and a primer he had produced (K. Holzknacht 1978:i-ii). The late Rev. K. Holzknacht reopened the Lutheran mission station in 1947, and served there for many years. He is still well respected within the Adzera community for his extensive knowledge of their language and customs. In addition to translated religious materials, Holzknacht wrote several articles on phonology and morphology (K. Holzknacht 1973a, b, and c), and several unpublished Adzera dictionaries (K. Holzknacht 1978). More recently S. Holzknacht (daughter-in-law of K. Holzknacht) published a grammar of the

Amari dialect of Adzera (1986). A few years later she published a comparative study of the fifteen languages of the Markham subfamily, including Adzera, in an attempt to reconstruct the parent languages (S. Holzkecht 1989). She has also published articles on word taboo in Adzera and other Markham languages (1987, 1988). There have been several SIL members to work among the Adzera people in different roles over the years, including M. Johnson, who, in addition to helping to train Adzera speakers to produce a revision/retranslation of the New Testament, has written some unpublished papers on the language (including Johnson 1997).

In addition to literature specifically on the Adzera language, others have written about Adzera culture, which cannot be completely separated from the language. K.E. Read published several articles about the Ngaruwapum district group, a political subgroup of the Adzera people (Read 1946, 1947, 1949). H. Holzkecht, son of K. Holzkecht, has also written about the Adzera and other Markham people from an anthropological perspective, as one who grew up in the culture (H. Holzkecht 1974, 1976).

1.4 Adzera Culture and the Place of Oral Narrative

Most Adzera speakers live in small villages scattered across the broad, grassy plains of the Markham Valley. Typically these villages are located near rivers or streams that feed into the wide and muddy Markham River. The people are primarily engaged in subsistence agriculture. Meat is not a daily part of their diet, though they do raise chickens and pigs for that purpose. They cook and live almost entirely outside, and only sleep and store possessions in their houses, which are typically square structures with pointed grass roofs and woven bamboo walls, built about five feet off the ground.

Before European contact, warfare and cannibalism between the Adzeras and their neighbors, and even between different Adzera groups, was common. People from opposing “district groups” frequently raided each other’s villages, and usually brought back slain enemy warriors to eat (Read 1946, 1949). This ceased after the introduction of Christianity by German Lutheran missionaries in the early 1900s, at least in the areas under Christian influence. Today it is a practice found only in the stories of times past.

The *yauŋ*, or shade tree,² is central to Adzera culture. This is where people interact as they seek refuge from the tropical heat. It is also where hosts demonstrate hospitality toward their (usually unexpected) guests, by welcoming them, inviting them to sit down, and offering them food. By doing so, these hosts gain status as *garam tsira* ‘big men’ (i.e. leaders or men of status and prestige in their communities). When “big men” are gathered in the *yauŋ*, it is not a place for women and children, who are expected to keep to the outskirts or find shade somewhere else to let the men discuss their business without interruption, except to bring food and tea. In such a situation it is the job of the women to cook and bring food to the *yauŋ*. It is the responsibility of the men of the household to be good and gracious hosts, by sitting and visiting with their guests and serving out the food to their guests first. In addition to certain gardening tasks, the men have the responsibility to arrive at community decisions by consensus, so they gather to discuss matters frequently. The women, on the other hand, have responsibility for the food preparation, cooking, washing, child care, cleaning of the *yauŋ* and areas around their houses, and most of the gardening.

² Actually, *yauŋ* means just ‘shade’ or ‘shadow’, but it most commonly refers to a shade tree. Often it accompanies words for trees: *gai yauŋ* ‘the shade of a tree’ or *nuwai yauŋ* ‘the shade of a mango tree’.

Since it is in the context of visiting that stories are usually told, storytelling is very common among the men. Women, to be sure, also tell stories to other women as they work together on special occasions or as they meet on the road, but their responsibilities prohibit them from doing so as often. Storytelling is a major source of entertainment within Adzera culture. It is one way to pass the time in a culture that has a slower pace (for men, that is) than our Western "rat race." The Adzera people cannot just turn on a television to be entertained whenever they want. Some do have shortwave radio, but that is of limited entertainment value. Oral stories are also a way to pass on traditions, cultural values, language, or even just everyday information. Even though there are some written materials in the language, it is still very much an oral society. For generations and generations information has been passed down orally. Adzera culture, history, and language are not written down in textbooks. Neither are there any Adzera novels. All of this is still orally communicated. So storytelling has a prominent role in Adzera society, and a prominent place under the shade tree. It does not matter if one tells a traditional story, a story about the way things were before, or a story about what happened yesterday on the way to the garden; all of these are fertile ground for oral narratives.

1.5 Data for the Current Study

I collected the data for the current study while living and working among the Adzera people as an SIL linguist/translator.³ My primary role is translation advisor and mentor for the Adzera mother-tongue translators, while my wife, Susan, is a literacy advisor. During our first two and a half years among the Adzera people (from April 1999 to November 2001),

³ My wife and I took over the SIL program from Morris and Wendy Johnson.

we were able to spend about 10 months in Binimamp village, where we live. There we have been “adopted” into the Bibuae family, on whose land we live.

Though I draw from other oral and written texts and elicited data, the current study is based primarily on two oral texts that I collected during our time there. I recorded both of them and transcribed and glossed them with the help of native speakers. Both texts come from older men—*garam tsira*²—from Binimamp village.

The first, which I have labeled text A, was given to me by Gadan Ayam. It is a story from his childhood about his father, who was an evangelist, and the opposition he received from an influential leader in the area where he served. This leader, whose name was Sinkom, threatened to kill and eat him, a vestige of the cannibalism of the past. The rest of the story describes the reactions to the threat, how the events unfolded, and the final outcome. Starting out as a third person account about the narrator’s father, the narrative switches to a first person perspective when the narrator himself becomes a main character as he accompanies his father on his mission. The text contains many quotations, which actually mark the peak of the narrative—the threat, a counter-threat from Sinkom’s own brother, the reactions of the main character and his son (the narrator), and a warning sent to them about the imminent danger. There are also a number of characters within the narrative, and frequent switching back and forth from one to another, especially at the peak of the story. The text of the narrative is found in appendix A.

The second story, text B, came from Wari Muruk. It is a story about his father-in-law, Tsinu, for whom the narrator always uses a pseudonym, Maramais or Maramai (the latter in this story), because of name taboos (cf. S. Holzknecht 1988). This narrative is about

Maramai going to find fish with his wife and finding a crocodile instead. There are many action sequences, including a flashback describing how the crocodile built her nest before Maramai happened upon it and found her. The story describes how Maramai speared the crocodile, and how he and his wife carried it back to the village, where it was divided, cooked, and eaten amid much celebration. There are fewer characters, fewer quotations, and longer action sequences in this text than in text A, and, therefore, much less switching between characters. The reader will find the full body of this text in appendix B.

I frequently use examples from these texts in the body of this thesis. When I do, their sentence number from the corresponding appendix appears between the example number and the data, for example: (12) B-8 *Da iyu git igi*. In this case, the data for example (12) comes from sentence #8 of text B, found in appendix B, and the reader can turn there for the full context. Examples without text numbers are from other texts or elicited examples. The speakers are bilingual in *Tok Pisin* (Melanesian Pidgin), the trade language, and occasionally mix in a word or two of it.

1.6 Expected Findings

The theoretical framework of this study and the predictions that that entails will be presented later in this thesis, so I do not go into much detail or give much explanation in this section. However, I present briefly some results that one would expect to find in the analysis of these texts. Throughout the texts there is frequent ellipsis of subjects, and sometimes objects as well. This ellipsis occurs in the absence of agreement morphology on the verb, making the identity of ellipsed referents dependent upon the preceding discourse. Therefore, one would expect these referents to be not only given, but easily identifiable based on an

immediately preceding mention, even more so than pronouns, which also mark highly salient, given entities. We should find that full noun phrases encode the entities that are the most difficult to identify (whether given or new), leaving pronouns somewhere in the middle. Givón (1983) predicts this ordering, as well as the following expectations, which are also fairly intuitive. Animate entities should be more topical and continuous than inanimate entities, and main characters should be more topical and continuous than other characters. In terms of distribution, more new entities should appear in non-subject positions than as subjects, which should be more topical or continuous than non-subjects. Furthermore, I would expect relatively few inanimate entities to occur as subjects, because Adzera lacks a passive construction.

1.7 Organization of the Thesis

The organization of the remainder of this thesis is as follows. Chapter 2 gives an overview of Adzera grammar in order to provide necessary background for the reader. Chapter 3 is a review of the literature on information structure theory that is relevant to this type of approach. Chapter 4 is an application of Prince's (1981) taxonomy of assumed familiarity to the referents in the Adzera texts. Chapter 5 applies to the Adzera texts Givón's quantitative approach to topic continuity. Chapter 6 summarizes the findings and conclusions of this study, and suggests areas for further study.

CHAPTER 2

AN OVERVIEW OF ADZERA GRAMMAR

2.1 Introduction

In this chapter I give a basic overview of certain aspects of Adzera grammar in order to help the reader understand the present study. The variety of Adzera in focus here is the central dialect chain, in particular, that which is spoken in the Sangang (or Sangan) area where I live and work. As this paper is primarily a discourse study, it is not my intention to present a thorough treatment of grammar here. Rather this covers some of the most basic aspects, as well as any which have particular relevance to the discourse study in subsequent chapters. For a more complete treatment of Adzera morphology and grammar, I refer the reader to S. Holzknrecht 1986, a description of the Amari dialect of Adzera, which, for the most part, applies to the central dialect chain as well. Holzknrecht points out the specific features of Amari that do not occur in other dialects or vice versa. I will, however, point out a few instances in which my analysis of the central dialect chain of Adzera disagrees with hers and others prior to that.

The chapter is organized in the following way. Before getting into syntax, we will look briefly at the phonemes and graphemes of Adzera. Then we will look at word order, which is the primary means for distinguishing subjects and objects. A description of noun phrase constituents follows, including a discussion of compounds and relative clauses.

Finally we will look briefly at verb phrases, including verbal morphology, subject and object ellipsis, serial verbs, and clausal complements.

2.2 Phonemes and Graphemes

This section does not include a phonological analysis of Adzera, with contrastive pairs as evidence. It simply presents the phonemes and graphemes of the language as an aid to the reader. For a more complete phonological analysis see S. Holzknrecht 1986 (cf. K. Holzknrecht 1973a). The phonemes presented here most closely reflect those proposed by S. Holzknrecht, though I note two exceptions. I include the phoneme /h/, even though it has likely been borrowed, as it is found in only one Adzera word *hai* 'yes'.¹ The other difference is that I exclude the phoneme /^hdʒ/, which I have not found in the language.² See table 1 below for the list of Adzera consonants. The prenasalized voiceless stops included in the table are prenasalized only after a vowel; initially or after a consonant, the nasalization is not pronounced. The syllable structure of Adzera suggests that prenasalized stops are single phonemes, not a cluster of two separate consonants, since the largest allowable consonant cluster is two, and the second consonant of a cluster can only be /r/ (S. Holzknrecht 1986:93).

¹ This also means 'yes' in Japanese (Goris and Okubo (eds.) 1995:583), and possibly could be the result of influence during the Japanese occupation of the area during World War II. Johnson (1997) also includes /h/.

² Evidence for the phoneme /^hdʒ/ is scarce. I have only seen a few questionable examples (K. Holzknrecht 1978:162, 344; repeated in S. Holzknrecht 1986:88; Johnson 1997:2), which I checked with two native speakers (one who spoke the Sangang variety of the central dialect chain and one from the Yarus dialect who is living in the Guruf dialect area). In every case except one they insisted the word had a different phoneme in it, either /^hts/ or /dʒ/. In the other instance neither was familiar with the word. Furthermore, no other voiced plosives (/b, d, g/) are prenasalized in Adzera, only the voiceless ones. Symmetry argues for a complete absence of prenasalized voiced stops. Wampar and Wampur, two closely related languages, also have voiceless but not voiced prenasalized plosives (S. Holzknrecht 1989:52-4). This is not to say that there cannot be such a phoneme in Adzera, perhaps even borrowed from a neighboring language, but there is insufficient evidence for it in the central dialect for me to include it with the other Adzera phonemes. K. Holzknrecht attempts to show examples of /mb/, /nd/, and /ŋg/, but the examples for the latter two are across morpheme boundaries, and I am unfamiliar with the third example (1973a:5).

Syllable patterns in Adzera include V, VC, CV, CVC, CCV, and CCVC (Johnson 1997:4; see also K. Holzknrecht 1973a:7).³

Table 1. Adzera Consonants

Graphemes	Phoneme	Phone(s)	Description
<P, p>	/p/	[p ^h], [p]	voiceless bilabial stop
<T, t>	/t/	[t ^h], [t]	voiceless alveolar stop
<Ts, ts>	/ts/	[tʃ], [ts]	voiceless alveo-palatal affricated stop
<K, k>	/k/	[k ^h], [k]	voiceless velar stop
<ʔ>	/ʔ/	[ʔ]	voiceless glottal stop
<Mp, mp>	/ ^m p/	[^m p]	prenasalized voiceless bilabial stop
<Nt, nt>	/ ⁿ t/	[ⁿ t]	prenasalized voiceless alveolar stop
<Nts, nts>	/ ⁿ ts/	[ⁿ tʃ], [ⁿ ts]	prenasal. voiceless alveo-palatal affricated stop
<ŋk, ŋk>	/ ^ŋ k/	[^ŋ k]	prenasalized voiceless velar stop
<ŋʔ>	/ ^ŋ ʔ/	[^ŋ ʔ]	prenasalized voiceless glottal stop
<B, b>	/b/	[b]	voiced bilabial stop
<D, d>	/d/	[d]	voiced alveolar stop
<Dz, dz>	/dʒ/	[dʒ]	voiced alveo-palatal affricated stop
<G, g>	/g/	[g]	voiced velar stop
<F, f>	/f/	[f]	voiceless labio-dental fricative
<S, s>	/s/	[s]	voiceless alveolar fricative
<H, h>	/h/	[h]	voiceless laryngeal fricative
<M, m>	/m/	[m]	voiced bilabial nasal
<N, n>	/n/	[n]	voiced alveolar nasal
<ŋ, ŋ>	/ŋ/	[ŋ]	voiced velar nasal
<R, r>	/r/	[r]	voiced alveolar flap
<W, w>	/w/	[w]	voiced labiovelar approximant
<Y, y>	/y/	[j]	voiced palatal approximant

Adzera has three contrastive vowels: /i/, /a/, /u/. K. Holzknrecht includes a fourth vowel /o/, originally posited as the only high back vowel by the Rev. Fritz Oertel, who worked as a missionary to the Adzeras prior to World War II (K. Holzknrecht 1978:i). Holzknrecht added /u/ and claimed to find minimal pairs distinguishing it from /o/ (1973a:4;

³ Here V represents a vowel (rather than a verb, as indicated on page xv), and C represents a consonant.

1978:ii). Like S. Holzkecht (1986) and Johnson (1997) I find no separate phoneme /o/, and the Adzera speakers whom I know do not recognize it as a separate phoneme. Words that K. Holzkecht spelled with an <o> they spell with a <u>. I have not heard an [o] in the language (except in borrowed words), and while some phonetic variation is possible, even likely with some speakers, [o] should be treated as an allophone of /u/. The phoneme /i/ is realized as [i] or [ɪ] or somewhere in between; /a/ is realized as [ɑ] or sometimes as [ə] when unstressed; and /u/ is realized as [u] or as a range including [ʊ] and possibly even [o]. In addition there are two diphthongs, /ai/ and /au/.

The reader may notice in the texts (included in the appendices) that certain letters not presented here, such as <j>, are used in proper names. Adzera follows the convention of using <j> rather than <dz> for people and place names that are spelled with <j> in English and *Tok Pisin* (Melanesian Pidgin). Furthermore, a previous orthography of Adzera (based on German and Jabêm/Yabêm) used <j> for the approximant [j], which is now represented as <y>. Consequently, names such as *Jisu* (Jesus) also use <j> to represent the /y/ phoneme. So <j> can be used to represent either /dʒ/ or /y/ in some proper nouns. Borrowed words with other letters not properly found in the Adzera alphabet would also employ extra letters, such as <o>, <e>, <ɻ> or <z>.

2.3 Word Order

Adzera follows a fairly rigid SVO word order, and objects never separate subjects and verbs. Example (1), below, shows two clearly transitive clauses with SVO word order:

and, as a result, must immediately precede the verb, which is why sentence c. above is not acceptable (if not ungrammatical) to Adzera speakers. So SV order is maintained.

Why is understanding word order so important? It is important because, in most cases, it is the only way to distinguish between subject and object. Adzera lacks the regular case markings that some languages have. Instead, the language employs SVO word order to mark its constituents, occasionally allowing an object to be moved forward as long as SV order is maintained.

2.4 Noun Phrases

2.4.1 NP Constituents

Noun phrases (NPs) will be very important to our discussion of forms of reference, especially in chapter 4. Adzera is a head-first language, so a head noun precedes all other NP constituents, with the exception of a possessor, which can be either a pronoun or an embedded NP. NP constituents can include descriptive modifiers, quantifier phrases, relative clauses, and demonstratives/determiners, roughly in that order. Descriptive modifiers can include other nouns, adjective phrases, or nominalized verb phrases.

Many concepts that English speakers express with adjectives, Adzera speakers express with verbs, such as hot or cold, bad, straight. These are called stative verbs,⁵ because they express a state of being rather than an action. The nominalized or “participle” forms of stative verbs can modify nouns just like true adjectives, which are also found in Adzera.

Compare (8a), a noun phrase, with (8b), a full sentence:

⁵ Cook 1998:8-9 gives a historical review of the treatment of both verbs and adjectives as predicates occupying the same general grammatical category, a position that is attributed to Lakoff and Postal (Lakoff 1966, 1970). See also Cruse 1986:217-8 and Comrie 1976.

- (8) a. *mur mais-an*
 snake bad-NMZ
 'a bad/poisonous snake' (or idiomatically 'the devil')
- b. *Mur i-mais.*
 snake R-bad
 'The snake is bad.' or 'Snakes are bad.'

Active verbs can modify nouns in relative clauses, as we will see in section 2.4.5.

2.4.2 Definite and Indefinite NPs

An indefinite determiner can follow the head noun and other modifiers to form an indefinite noun phrase. Speakers usually use the determiner *maɲan* 'a/one/another',⁶ as in the following examples:

- (9) B-11 *fugai maɲan*
 crocodile a
 'a crocodile'
- (10) A-29 *garam tsira' maɲan*
 man big a/one.of
 'one of the big men (i.e. leaders)
 (sometimes also 'a big/large man' or 'a man of high status')

The NP in example (10) can be ambiguous. In the present context it means one of a given set, namely the "big men" or leaders that every village has (not just formally but in terms of community status), but the phrase could simply refer to any man of high status or, less commonly, a large man. Sometimes indefinite noun phrases contain only the head noun, as in example (11a). Note, however, that a head noun by itself can also be definite, as in (11b).

- (11) a. A-5 *Da waʔ, dzarʔ iwaʔ.*
 and (um) pig R-come.out
 'And a pig came out.'

⁶ The indefinite determiner *maɲan* 'a/one' is different than the numeral one *bitsintaʔ* or *bits*.

- b. A-6 *Ini ntaŋa dzar?*
 R-want spear-NMZ pig
 'He wanted to spear the pig'

It is common, however, to form definite noun phrases by adding a demonstrative determiner after the head noun, as in example (12) below, which immediately follows the introduction of an indefinite noun phrase consisting of the head noun alone, i.e. *git* 'a nest' (sentence B-7).

- (12) B-8 *Da iyu git igi...*
 and R-take nest that
 'She took the nest...'

In this example *igi* 'that/there', which is the default distal demonstrative, carries the sense of the definite article, pointing back to a given antecedent rather than to a spatial relationship within the speech situation.⁷ Table 2 shows Adzera demonstratives:

Table 2. Demonstratives

Dem.	Meaning	Dem. w/ <i>aru</i>	Meaning
<i>(n)ani</i>	here/this (very close to speaker)	<i>aruani/arani</i>	(this) right here
<i>(n)igi</i>	there/that (may be close to hearer)	<i>arunigi/arigi</i>	(that) right there
<i>(n)aga</i>	there/that (far away, in or out of sight)	<i>arunaga/araga</i>	(that) right over there
<i>(n)ugu</i>	previously present or mentioned	<i>arunugu/arugu</i>	that one previously present

Sometimes the morpheme *aru* precedes a demonstrative to make the reference more precise: *arunigi* 'right there'. These precise forms, which can be contracted (e.g. from *arunigi* to *arigi*), also can function as demonstrative pronouns, as in example (13):

- (13) A-50 *Da isalim arigi*
 and R-send that.one(just mentioned)
 'And he sent him'

⁷ For a discussion of discourse vs. spatial deixis, see Fillmore 1997.

The shorter forms, when used as demonstrative pronouns, are preceded by <n>, making *nani* ‘this here’, *nigi* ‘that there’, etc., as indicated in the table. Demonstratives also express time, since time is considered in spatial terms (cf. Fillmore 1997). Consider the examples in (14):

- (14) a. (*gubu?*) *ar(u)ani*
 day right.here/now
 ‘today’ or ‘now’ (cf. example (1) above)
- b. *sunta? maŋan ugu*
 Sunday a before
 ‘a week ago’
- c. *gubu? buntun aga*
 day end then(future)
 ‘the last day’ (prophetic)

2.4.3 Possession

There are two ways to show possession in Adzera. One is by means of possessor agreement marking. All nouns that indicate a possessor must have possessor agreement. The other way to show possession is via an overt possessor preceding the noun, which can be a pronoun, a proper noun, or a common noun phrase. This can be omitted when the possessor is clear, leaving only the agreement marking.

There are some Adzera nouns that must always have a possessor, such as body parts and kinship terms. These are called inalienably possessed nouns, because the noun cannot be separated from its possessor.⁸ The roots of inalienably possessed nouns cannot occur in isolation, but must take a person agreement suffix, either first (*-ŋ?*), second (*-m*), or third (*-n*) person, as example (15) shows. The generic (default) form is the third person.

⁸ See Velazquez-Castillo 1996:23ff for a historical review (1910-1996) of the literature on and terms applied to the alienable-inalienable distinction.

- (15) a. *baŋi-ŋʔ* 'my/our hand(s)'
 b. *baŋi-m* 'your hand(s)'
 c. *baŋi-n* 'his/her/its/their hand(s)'

These are the short forms, which are commonly used as vocatives or as shortcuts when emphasis is not required. There are long forms, which can be used for emphasis or in any way that the short forms are used, even as vocatives that emphasize the relationship. The long forms add the possessive morpheme *-ga*, which also must be specified for person in the same way, as example (16) shows:

- (16) a. *baŋi-ŋʔ-ga-ŋʔ* 'my/our hand(s)'
 b. *baŋi-m-ga-m* 'your hand(s)'
 c. *baŋi-n-ga-n* 'his/her/its/their hand(s)'

The possessor-marking nasals in the middle of the words in (14) assimilate to the <g> in the central dialect of Adzera, and in most other dialects (S. Holzknecht 1986:98), so that *baŋingam* becomes *baŋiŋgam* and *baŋingan* becomes *baŋiŋgan*.⁹ In the Sangang area, where I work, people frequently contract the first person suffix *-gaŋʔ* to *-gaʔ*, and often nasalization is heard on the vowel. Possessor marking on all other (i.e. alienably possessed) nouns follows the form: ROOT + {*-gaŋʔ*, *-gam*, *-gan*}, except that the /g/ is dropped after roots ending in consonants other than /ŋ/, /ŋ/, and /n/, and occasionally /m/.¹⁰

⁹ Many speakers do not pronounce the glottal stop in the middle of *baŋiŋʔgaŋʔ*, so I do not include it in the spelling: *baŋiŋgaŋʔ*.

¹⁰ Hence, *gumaŋʔ* 'my work' (from *gum* + *aŋʔ*) and *naŋgaŋʔ* 'my talk' (from *nan* + *gaŋʔ*), but *naŋgaŋʔ* 'my thing(s)/food' (from *nam* + *gaŋʔ*). Thus 'my talk' is indistinguishable phonetically from 'my thing(s)/food'. But I understand that in some speech communities *namaŋʔ* is acceptable.

So far we have looked at only possessor agreement marking, which can stand alone when the possessor is clear. Now we will look at examples of overt possessors. Example (17) shows a proper noun as an overt possessor:

- (17) A-8 *Augustin baji-n*
 Augustin hand-3pss
 'Augustin's hand'

The next two examples show full NPs as possessors. In example (18) a proper noun is the possessor of a full NP, which, in turn, is the possessor of another NP.

- (18) A-49 [*Gurump naru-ŋ-gan*] *bijaŋ-gan*
 Gurump child-3pss name-3pss
 'Gurump's son's name'

- (19) A-29 [*garam tsira' manjan*] *bijaŋ-gan*
 man big a/one.of name-3pss
 'the name of one of the big men'

Personal pronouns frequently precede possessed nouns as possessive determiners, but the only information they provide that is not included in the agreement marking is number; they distinguish between singular and plural, as shown in example (20):

- (20) B-10 *dzi bu-ŋ-ga' Maramai'*
 1sg inlaw-1pss Maramai
 'my father-in-law, Maramai'

Finally, the possessive morphemes, *gaŋ'*, *gam*, and *gan*, can function as possessive determiners,¹¹ but they are not marked for singular or plural. The most common is the third person pronoun *gan*, shown in example (21):

- (21) B-3 *gan yafas-an*
 3pss fish-3pss
 'their fish' (but also 'his/her/its fish')

¹¹ Combined with personal pronouns, these possessive morphemes form possessive pronouns: *dzi gaŋ'* 'mine'.

2.4.4 Pronouns

I have briefly discussed the use of personal pronouns within noun phrases as possessive determiners, and, of course, pronouns can substitute for noun phrases, but I have not introduced the pronominal system. What follows in table 3 is a very basic outline of the personal pronouns in Adzera. Once again I refer the interested reader to S. Holzknecht 1986 for a more thorough treatment.

Table 3. Personal Pronouns

	SINGULAR		PLURAL	
	Subject/Possessor	Subject/Object	Subject/Possessor	Subject/Object
1 st Person Exclusive	<i>dzi</i>		<i>aga</i>	<i>agai</i>
1 st Person Inclusive			<i>agi</i>	
2 nd Person	<i>u</i>	<i>agu</i>	<i>agam</i>	
3 rd Person	<i>aranjan</i>		<i>rib</i> + DEM.	
			<i>rib</i> + <i>aranjan</i> (Central dialect)	
			<i>tuas</i> + <i>aranjan</i> (Sangang area)	

The reader will notice that Adzera does not make a distinction between masculine and feminine third person singular pronouns. Nor do pronouns distinguish between subject and object in most cases, but there are a couple of instances where they do. The pronoun *u* ‘you (sg)’ cannot be an object, and while *agu* ‘you (sg)’ may be a subject, especially when emphasis is desired, it is most common as an object or a vocative, as in example (22):

- (22) *Agu, u sarj' u fada Lae da tata??*
 2sg 2sg can 2sg go-NMZ Lae TIME tomorrow
 ‘Can you go to Lae tomorrow?’ (i.e. ‘You should go to Lae tomorrow.’)

Likewise, among plural pronouns, *aga* ‘we (exclusive)’ cannot be an object, while its counterpart *agai* ‘we (exclusive)’ can be either a subject or object, but is more common as an

object. The third person singular pronoun *arajan*, and third person plural pronominal forms containing it, sometimes contract in subject position by dropping the [n]. All other pronouns make no subject-object distinction. Again, this is why word order is important in Adzera.

There is no single third person plural pronoun in Adzera. Rather there are four traditional forms constructed by joining the third person plural morpheme *rib* with the demonstratives *ani*, *igi*, *aga*, and *ugu*, to indicate ‘those here’, ‘those there (near)’, ‘those there (far)’, and ‘those that were here before’, respectively. *Ribigi* is the general default form for ‘they’, which is usually used in narrative to refer to a party that has been mentioned. In the central dialect chain, *ribarajan* ‘they’ is a newer, commonly used generic form, and in the Sangang speech community, *tuasarahan* ‘they’ is frequently used, and seems to be the preferred form for referring to groups previously mentioned. The morpheme *rib* also occurs in conjunction with place or people names or descriptions to mean ‘the people/those of X’, for example, *rib Dzigintsuan* ‘those of Dzigintsuan’ or *rib mamai* ‘the mountain people’. However, these are really full noun phrases because of their descriptive nature, and of course, that is how we must translate them. We will deal with these forms more in chapter 4.

2.4.5 Relative Clauses

Relative clauses in Adzera are not marked by relative pronouns as in English, but instead employ nominalized verb phrases. The result is that it can be difficult to distinguish Adzera relative clauses from other types of clauses that use the same kind of verbal construction. I will give an outline of relative clauses here so the reader can see a few clear

examples, and I will discuss ambiguous cases that have relevance to chapter 4 in section 2.5.4.

S. Holzknrecht claims that, for the Amari dialect at least, the head noun modified by a relative clause is optionally followed by the demonstrative *ugu*, and then by the relative clause, which always ends with the demonstrative *ugu* (S. Holzknrecht 1986:111). She gives seven examples of this construction. The problem, however, is that in each case the action expressed within the relative clause is in past tense, occurring at a time prior to the action in the main clause. I would argue that Holzknrecht's treatment of relative clauses is too narrow because she focuses on a limited set of past tense circumstances, eliciting the use of the demonstrative *ugu*.

Relative clauses are not always as clear-cut as Holzknrecht (1986) indicates. I have found much more variation in the Sangang speech community, including the use of other demonstratives, like *igi*, as in example (23) below, or the use of no demonstratives at all, as in (24):

- (23) ...*da rab-a [ntsim [agam u-dan]_{RC} igi]_{NP...}*
 and chop-NMZ betelnut 2pl take-NMZ that
 '...and cutting the betelnut (that) you took...' (within a subordinate clause)

- (24) [*Nan mais-an idzuwai santan [agam narj-a da rungam*
 talk bad-NMZ what all 2pl make-NMZ to each.other
da aga tias-a rai]_{RC}]_{NP} da Nifutsarif itias rai ibinigi
 and 1pl.ex wipe-NMZ away and Lord R-wipe away R-like-that

'Whatever bad talk [(that) you have made to each other and we have forgiven], the Lord has forgiven also.'

Sometimes Adzera speakers use the morpheme *aru* to introduce a relative clause (cf. section 2.4.2 above). In this type of relative clause construction the descriptive clause is sandwiched

between *aru* and a demonstrative, usually *igi*: HEAD NOUN *aru* [DESCRIPTIVE CLAUSE] *igi*.

Consider example (25) from text A:

- (25) A-77 *Da Sinkom ini [nanggan [aru a?, is-a, tsarif-a ruan*
and Sinkom R-say talk-3pss SPEC (uh, kill-NMZ) exalt-NMZ self-3pss

*i is-a ramay?]*_{RC} *igi]*_{NP}
about kill-NMZ father-1pss that

‘Sinkom spoke his boastful threat about killing my father’
(or ‘his words, the ones with which he exalted himself about killing my father’)

For lack of a better term, I call *aru* a specifier (SPEC), but it appears to fit the mold of a relativizer. It is the closest thing to a relativizer that I have found in Adzera, and it appears at the beginning of a relative clause between the head noun and the rest of the clause, which is where one would expect to find a relativizer (Bickford 1998:339). I give more examples of relative clauses in appendix C,¹² including those that have *aru*, those with various demonstratives, and those without either. The point is that there is much variation in the construction of relative clauses, which leaves the door open for confusion with other constructions.

2.5 Verb Phrases

There is much that could be said about verb phrases, but space does not permit us to discuss all of it here. Rather I will only present here the aspects that are most relevant to our discussion in chapters 4 and 5. In particular, we will take a brief look at morphology—what

¹² Also in appendix C, the reader will find examples of compounds that are very similar to relative clauses, which are formed by a head noun plus a nominalized verb phrase that indicates a person’s profession or “personal idiosyncrasies and characteristics” (S. Holzknecht 1986:109).

is and what is not encoded by the verb. Then we will look at ellipsis, serial verbs, and, finally, clausal complements, and how they are distinguished from relative clauses.

2.5.1 Verbal Morphology

We have seen in the section on noun phrases that nominalized verb forms can modify nouns. This is the form that is equivalent to English participles, gerunds, and infinitives, except that Adzera has only one form to cover all of these uses. It is also used to form certain nouns, such as ‘food’, which is literally ‘thing to eat’: *nam ga-dan*. As the previous examples show, the nominalized form consists of the verb root plus the suffix $-(d)a(n)$, the actual form of the suffix depending on the phonological environment that precedes and follows it. This is the only suffix that Adzera verbs can take. However, verbs can take prefixes that indicate mood, tense, and aspect, as well as freestanding preverbal morphemes. I present a simplified summary of the prefixes and preverbal morphemes in table 4.

Table 4. Preverbal Morphemes and Prefixes

Preverbal Morpheme	Prefix
<i>anujʔ</i> negative	<i>i-</i> realis
<i>buɟ/buɟʔ</i> present perfect aspect (or <i>bu-</i> + ROOT)	
<i>ruɟ/ruɟʔ</i> continuous aspect (or <i>ru-</i> + ROOT)	
<i>buɟʔ</i> future tense	<i>a-</i> irrealis ¹³
<i>ru-</i> continuous aspect	<i>w-a-</i> imperative (negative) <i>n-a-</i> hortative (future) <i>m-a-</i> potential/uncertain <i>d-a-</i> contrary-to-fact

¹³ K. Holzkecht (1973c) and S. Holzkecht (1986) treat this morpheme as part of the future tense morpheme, but for reasons that I will not explain here, I treat them as separate morphemes. They also treat all the morphemes in this chart as prefixes.

Adzera makes a distinction between actions or conditions that are already realized (realis) and those which are yet unrealized (irrealis). Realis mood marks those actions or states that happened or were in effect in the past or are happening or are in effect in the present. Irrealis mood marks those actions or states that will or should happen or be in effect in the future. The realis prefix *i-* is the most frequently used. Verbs carrying this prefix can be either in present or past tense, so other markers or context must disambiguate the two. Another thing to note is that Adzera has no passive voice morpheme or any other strategy to passivize verbs.

The most important thing for the reader to notice in table 4 is the complete absence of any kind of person, number, or gender agreement morphology on the verb. Verbal morphology reflects only mood, aspect, tense, negation, and nominalization. This is highly relevant to our discussion, since subjects (and, less commonly, objects) can be omitted. Sometimes the realis prefix is dropped or unheard immediately after certain pronouns, but the presence or lack of the prefix does not relate to agreement; it is a phonological process.

2.5.2 Ellipsis of Subject and Object

Even though verbs carry no subject or object agreement, speakers often omit subjects and objects. I refer to this kind of omission as ellipsis (or zero anaphora). Ellipsis of subject is extremely common in Adzera narrative; in fact, it is the norm when the identity of the subject is known. Some of the example sentences we have already seen contained ellipsis, examples (4), (12), and (13), for instance. Consider the clauses in example (26), which immediately follow the introduction of their subject, a crocodile:

- (26) B-6 *Da Ø inaj rujum tsiraʔ, Ø ipanu rujum tsiraʔ ifa fa waʔ,*
 and R-make deep.spot big R-dig deep.spot big R-go go (um)
da Ø igirʔ gin da Ø iwanun.
 and R-lay 3obj and R-pregnant
 ‘(She) dug a very deep hole; (she) dug and dug until it was very
 deep, and (she) lay in it and was pregnant.’

Example (26) shows the first four clauses of a string of 16 clauses without a single overt subject. The implied subject, the crocodile, does not change, and is only overt in its first mention as a direct object. Going back to the last known referent is the only way to identify the agent of the verbs. In this case, the referent could be the ellipted subject referent in the preceding sentence (i.e. the man, Maramai), or it could be the crocodile, who is the object of the same sentence. But the descriptions soon clear up any confusion in the hearer’s mind—no man can be pregnant nor would a man normally dig a deep hole in the water and lay in it, especially after seeing a crocodile in the water. So communication continues even in the absence of overt subjects.

The next sentence contains not only ellipted subjects but an ellipted object as well:

- (27) B-21 *Ø Ifinti sijan igi sib da Ø imunti da Ø intajʔ Ø.*
 R-charm spear that already and R-stand and R-spear
 ‘When (he) had performed the ritual blessing on the spear, (he) stood
 and speared (the crocodile).’

In the context of this sentence it is perfectly clear who is spearing what, so both can be omitted without causing miscommunication. Ellipsis is a key topic in chapter 5, as I compare the continuity of ellipted entities with that of entities encoded by other forms of reference.

2.5.3 Serial Verbs

Adzera frequently employs sequences of verbs, which, together, communicate one overall event or action. These are known as serial verbs, because they occur in series.¹⁴ Verbs of motion and/or direction, such as ‘go’, ‘come’, ‘(go) up’, and ‘(go) down’, are common in serial constructions. Often one or more verbs of a sequence have an object. A common sequence is: go-get-come (see (28d) below). In example (28) I have numbered each verb in a series with a subscript numeral to highlight the verb sequences.

- (28) B-9
- a. *Iyu₁ git igi iyab₂*
R-get nest that R-go.up
 - b. *ya yari*
and R-build
 - c. *ya irua[?]₁ iru[?]₂ ifan₃*
and R-fall R-go.down R-go
 - d. *ya itip₁ ifa₂ yu₃ a[?] puatsi iba₄*
and R-repeat R-go R-get (uh) grass R-come
 - e. *ya iba₁ iyab₂*
and R-come R-go.up
 - f. *ya irim*
and R-put
 - g. *ya irua[?]₁ iru[?]₂ ba₃.*
and R-fall R-go.down (R-)come

‘(She) took the nest up and built it, and (she) went down and brought back some more grass and came back up and put (it on the nest) and came back down.’

This string of seven clauses about a crocodile building her nest is a beautiful example of both ellipsis, which we just examined, and serial verbs. There are serial verbs in a., c., d., e., and

¹⁴ For a more complete discussion of serial verbs, see Lefebvre (ed.) 1991, particularly Givón in that volume.

g. Looking at each sequence, one can see how it acts more like one verb than several: a. 'took [the nest] up'; c. 'went down'; d. 'went and brought back more [grass]'; e. 'came up'; and g. 'came down'. These are by no means the only serial verbs in this text; on the contrary, there are so many that I had trouble deciding which to show, and the other text has many as well. However, the verbs in this example should suffice to give the reader an adequate understanding of what serial verbs in Adzera look like and how they function. I discuss how I count the subjects of serial verbs in chapter 5.

2.5.4 Clausal Complements

The construction of clausal complements is relevant to our discussion in chapter 4 because distinguishing them from relative clauses is not always easy, and classifying a construction alternately as a relative clause or a complement clause can affect the classification of referents within that construction. Before we look at examples of clausal complements (and ambiguous examples), we need to look at a few basic forms.

In Adzera there is a freestanding multi-purpose morpheme *i* that can be difficult to gloss. It can introduce cause or purpose clauses; it can function like a preposition such as 'about', 'for', or 'with' (instrument); and it can precede the object or complement of certain verbs, whether the latter is a phrase or a clause. Some of these verbs require either *i* or the third person object pronoun (*gin*), which S. Holzkecht analyzes as a form of *i*, suggesting it could be analyzed as *i* + *-n* for third person (1986:131). As she points out, (*g*)*in* can mean not only 'him', 'her', 'it', or 'them', but also 'with it' (instrument), 'about it', 'for them', etc.

(Holzknecht 1986:131). Verbs such as ‘know’, ‘ask’, ‘hear’, and many others require this object pronoun or *i* plus an object. Consider the following examples in (29):

- (29) a. *riɣant-a gin*
hear-NMZ 3obj
‘to hear something’
- b. *Wa-riɣant in.*
IMP-hear 3obj
‘Listen (to it).’
- c. *Dzi i-riɣant i kriɣaŋkriɣ.*
1sg R-hear Obj bell
‘I hear the bell.’
- d. (B-16) *Da fugai i-riɣant i riririɣ ba-dan*
and crocodile R-hear Obj/Comp noise come-NMZ
‘The crocodile heard a/the noise coming’
(i.e. of Maramai coming near to look)
- e. (A-68) *Raiyi, Sip i-riɣaŋ’ i aga mpa-da puatsi.*
later Sip R-hear Comp 1pl.ex stay-NMZ grass
‘Later, Sip heard that we were living in the bush.’

In (29) pronominal objects are marked by *gin/in*; in (29c) the object marker *i* precedes the object NP; while in (29e) a clausal complement replaces a simple object NP, so I call the *i* a complement marker. What is heard in (29e) is the news that they were living in the bush, which we must translate as a complement clause in English. It is tempting for English speakers to translate *i* in (29e) as ‘that’ and call it a complementizer. However, as we see in (29c), this verb requires the marker for any object, even the simplest noun phrases.

I have skipped over (29d) because its structure is ambiguous. What was heard was not news that a noise was coming, but the actual noise as it approached. Is this a clausal complement or an NP containing a VP that modifies the head noun, *riririɣ* ‘noise’, such as a

relative clause? If it is the latter, then one could translate it as ‘the noise that was coming’ or ‘the coming noise’ or perhaps even ‘the approaching disturbance’. This structure is even problematic in English. If one says, “I heard a noise approaching,” does that mean “I heard that a noise was approaching” (complement clause) or “I heard a noise that was approaching” (relative clause) or something else? This construction could be considered to be a reduced relative clause, or it could be a complement clause without a complementizer (see footnote 15 on page 34, below). On the other hand, if one says, “I heard something was approaching,” this must be a complement clause without the complementizer. Adzera, however, does not have “be” verb auxiliaries to differentiate the constructions. Again, the main obvious difference between (29d) and (29e) is what is heard. In d., the crocodile heard the actual noise. In e., Sip did not actually hear them living there, throwing wild parties or doing anything else; he just heard about it. Yet syntactically the construction in d. looks very similar (at least on the surface) to the complement clause of e. Is it purely a semantic difference or are these different clause types with an underlying structural difference?

Other verbs also take the object/complement marker *i* in some cases. Note its use in the next two examples:

- (30) A-51 *Wa-runt da wa-ni da mision*
 IMP-run and IMP-say to missionary
- i ya' Sinkom buŋ' a' na-bia' ya na-is a' mision*
 Comp. (uh) Sinkom (FUT ah) FUT-come and FUT-kill (uh) missionary
- da na-ga da gubu' arani.*
 and FUT-eat TIME day this.one
- ‘Run and tell the missionary that Sinkom will come and kill him and eat him today.’

- (31) *Dzi i-ni da Tengeng i papiar biŋanŋan sīb.*
 1sg R-say to Tengeng Obj. book name-3pss already
 'I already told Tengeng the name of the book.'

In both (30) and (31) the content of the verb *ni-da* 'to say/tell' follows the object/complement marker. It marks a complement clause in (30) and a noun phrase in (31). Notice also that in (30) none of the verbs within the complement clause are nominalized; all are in the future tense.

We have seen the use of the object/complement marker *i* with both complement clauses and simple objects with verbs that require it. Next I will show examples of complement clauses without the marker. This is common with the verb *tsaŋan-dan* 'to see', which, unlike *riŋant-a giŋ* 'to hear something', does not require the marker *i*.

- (32) A-17 *Da ramaŋ' i-tsaŋa [riŋ Dzigintsuan bugin-da wa' ramaŋ' i]*
 and father-1pss R-see 3pl Dzigintsuan dislike-NMZ (um) father-1pss
 'And my father saw that the people of Dzigintsuan did not want him'

In this case, it is clearly the whole proposition in brackets that the character "saw" or actually perceived. This is an unambiguous example of a complement clause without a complementizer or complement marker. The *i* at the end of the clause is a strengthening or confirming particle, like 'indeed' (K. Holzknacht 1973c:21). Example (33) contains a complement clause with the same verb in the matrix clause:

- (33) *Da araŋa i-tsaŋa [wap da gai [araŋa mpa-da giŋ]_{RC} igi su-da tsira?].*
 and 3sg R-see forest and tree 3sg sit-NMZ 3obj Dem. become-NMZ big
 'And he saw the forest and the tree where he sat become open space.'

Neither of the above two examples with the verb *tsaŋan-dan* 'to see' has a complementizer or any kind of complement marker, and all of the embedded verbs in them take the nominal

forms. But, as in (30) above, it is not always the case that the embedded verbs take the nominalized forms, which we see in example (34):

- (34) *da i-tsaŋa [[gai kuadaŋ nidzun dzub-an]_{NP} i-rua'*
 and R-see tree {species} seed/fruit ripe-NMZ R-fall
da [dzaŋ' bum-an]_{NP} i-gan].
 and pig wild-NMZ R-eat
 'and he saw ripe seeds/fruit of the *kuadang* tree fall(ing) down and wild pigs eat(ing) them.'

An alternative translation of example (34) might be: 'and he saw that ripe seeds/fruit of the *kuadang* tree had fallen/were falling down and that wild pigs had been eating them.'¹⁵

The realis verbs in (34) show that it must be a clausal complement, since this type of verb is sometimes allowed in complement clauses, but not in relative clauses.

To summarize our discussion, clausal complements are not always easy to distinguish from relative clauses. There are many ambiguous examples, and more research needs to be done to disambiguate the two constructions. Some verbs require the object/complement marker *i*, regardless of whether the object of the verb is a simple NP, an NP modified by a relative clause, or a complement clause. Other verbs do not take this marker, even when they have clausal complements of the type that we would introduce with *that* in English. The marker, then, gives us little help as we try to clear up the ambiguous examples. As noted

¹⁵ Precise translation of example (34) is difficult because it is unclear whether the person actually saw the seeds fall and the pigs eat them or if he just saw that they had been falling and that pigs had been eating them. In English the complementizer *that* indicates the latter meaning, but the absence of it indicates the former. So there is a semantic difference between the two forms, but perhaps not a structural one. If someone says, (A) "I saw the tower fall(ing)," it is not the same as saying, (B) "I saw that the tower fell," nor is it the same as (C) "I saw the tower that fell." (A) essentially means, "The tower fell and I saw it happen." It is the falling that was seen, instead of the result of the falling, as in (B). Both (A) and (B) contain clausal complements with two different semantic interpretations relating to aspect, while (C) contains a relative clause. This has implications for the previous examples as well, since one cannot assume that the inappropriateness of a translation employing the complementizer *that* necessitates treating the example as a relative clause. Thus one may do well to interpret the construction in (29d) as a complement clause, as in (34), but it is not entirely clear.

above, the classification of referents within such a construction depends on the identification of the construction itself, so I have attempted to lay out the issues clearly in this section.

CHAPTER 3

THE THEORETICAL FRAMEWORK FOR THE PRESENT STUDY

3.1 Introduction

One of the things that sets humans apart from animals is the ability to communicate not just emotions, warnings of danger, and information about their immediate context, but also complex information, ideas, and feelings about events or situations far removed from them in time and place. People employ rich vocabularies and grammars to communicate effectively with each other, sharing knowledge and ideas with each other. This is achieved by joining one concept to another, much like building blocks. New “blocks” of information—“information units” (Halliday 1967)—are stacked upon those that are already firmly established. Step by step a speaker takes a hearer from one point of consciousness to another, creating at each point in the course of a conversation an active awareness that did not exist the previous moment. New information is built upon old (given or established) information. The new information is not necessarily brand new to the hearer’s long-term memory, but it is certainly new to the present consciousness of the hearer, as we will discuss below. The organization of information within a discourse, which in turn determines constituent structure, has been called “information structure,” a term proposed by Halliday (1967:200). Much has been written about information structure, though not necessarily under that label. Below we will examine the views of some of the most prominent contributors in this field as a backdrop for the present study.

Languages have numerous strategies to smoothly handle the flow of information. Given information is not marked in the same way that new information is; therefore, a particular referent or idea will not be manifested in the exact same way each time it is used. If it were, then clauses would be too long and cumbersome to facilitate smooth and efficient transfer of ideas. Languages are organized so that speakers can refer to given information with a minimum expenditure of time and energy and can often leave the information implied. This economy is balanced with the need to be informative and not omit crucial information, all of which is expressed well by Grice's maxims below. The proper balance of these factors affects the forms of reference units of information take.

3.2 Grice's Maxims

Grice's contribution to this area of study outlines very well some of the universal rules of language that other contributors assume, which are basic to our understanding of the issues involved. In his 1967 lecture which was later published (Grice 1975), he formulates an overriding principle and several more specific maxims that participants in a conversation would be expected to follow. The main principle is what he calls the "Cooperative Principle," and he states it as follows: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (Grice 1975:45).

There are four categories in which this general principle of cooperation should manifest itself: quantity, quality, relation, and manner. Although there is much overlap between the categories of quantity, relation, and manner, our study is primarily concerned with the category of quantity, under which he gives two maxims: 1. "Make your contribution

as informative as is required (for the current purposes of the exchange),” and 2. “Do not make your contribution more informative than is required” (Grice 1975:45). These maxims help to determine the forms of reference people use in their speech. Speakers will avoid using forms that provide either too much or too little information than is necessary.

3.3 Given vs. New Information

3.3.1 Given and New Defined

Many authors have written on the subject of given (or old) versus new information (Halliday 1967, 1970; Chafe 1974, 1976, 1987, 1994; Haviland and Clark 1974; Clark and Haviland 1977; Prince 1981, 1992; Ward 1990, 1999; Birner 1994, 1997; Lambrecht 1994; *inter alia*), but not all of them have used these terms in the same way. There are at least three ways of defining the terms, and we will begin by looking at three different perspectives, namely those of Halliday, Chafe, and Haviland and Clark. Ultimately I will be using the approach of Prince (1981, 1992), who does a good job of summarizing and integrating these three perspectives of information status.

Halliday views given vs. new in terms of recoverability and predictability. Prince summarizes Halliday’s view nicely like this: “The speaker assumes that the hearer CAN PREDICT OR COULD HAVE PREDICTED that a PARTICULAR LINGUISTIC ITEM will or would occur in a particular position WITHIN A SENTENCE (1981:226, emphasis hers). Therefore, to Halliday, a contrasted referent represents new information, “not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse” (Halliday 1967:204). Again he states, “‘new’ is always to be interpreted as ‘contrastive’, as

contrary to some predicted or stated alternative” (Halliday 1967:206). To Halliday, then, given information is that which is expected in a certain context, and new information is unexpected, even if previously mentioned in the discourse.

In contrast, Chafe laments the “unfortunate tendency of both linguists and psychologists to pick foci of contrast as paradigm examples of new information” (1976:38), citing the primary examples of new information used by Halliday (1970) and Clark and Haviland (1977). By Chafe’s definitions the given-new distinction and contrastive focus are independent of one another, and “a contrastive referent may be given, accessible, or new” (1994:77), since the key issue for him is whether or not the referent is already in the consciousness of the hearer. Chafe (1974, 1976) defines the term *given* as information that the speaker assumes to be in the consciousness of the hearer at the time it is spoken, whether or not it is predictable. New information, on the other hand, is “what the speaker assumes he¹ is introducing into the addressee’s consciousness by what he says” (Chafe 1976:30). His view is what Prince (1981) considers givenness in terms of “saliency.”

Chafe suggests that the terms “already activated” and “newly activated” would more accurately capture his distinction between given and new (1976:30). In more recent work he proposes a categorization of given-new status in terms of three “activation states”: active, semi-active, and inactive (Chafe 1987:25). An active (or given) concept is one that is in the present consciousness of the hearer. An inactive (or new) concept is one that is not currently

¹ Because of the nature of this study and the frequent reference to hypothetical persons, such as “the speaker” and “the hearer,” and subsequent use of pronouns to refer to the same, I follow the somewhat dated convention—found here in Chafe’s quotation—of using masculine pronouns (*he, him, his*) in a generic sense to refer to non-specific gender-neutral persons. It is not my intention to exclude or devalue women in any way, but rather to avoid lengthy and awkward constructions caused by repeated use of the forms *he/she, him/her, and his/her*, which could potentially disrupt coherent communication. Therefore, where a specific person is not indicated, the reader should take *he, him, or his* to include women as well as men.

in the active consciousness of the hearer, though it may be stored in long-term memory. A semi-active—also called “accessible”—concept is “one that is in a person’s peripheral consciousness...but which is not being directly focused on” (Chafe 1987:25). Because this terminology avoids the ambiguity inherent with the terms *given* and *new*, others, too, have followed suit by employing some or all of these terms (cf. Gundel et al. 1993; Lambrecht 1994).

Chafe further discusses how the status, or activation state, of a referent can change throughout a discourse as it is activated or deactivated. He refers to these changes in terms of activation “cost.” To activate a concept requires effort, but it is considerably more difficult to activate an inactive concept than to activate a semi-active one. A concept can be deactivated from an active state to a semi-active state at no cost when it is no longer mentioned; it is automatic because of the limited capacity of one’s “focal consciousness” (Chafe 1987:31). According to Chafe, this is one of two ways a concept can achieve a semi-active state, the other being related to schemata, to be discussed in section 3.4.5 below.

Haviland and Clark (1974) have yet another way of defining *given* and *new*—in terms of “shared knowledge” (Prince 1981). They define old or *given* information as “what the listener is expected to know already” and new information as “what the listener is not expected to know already” (1974:512), rather than what is or is not active in the listener’s mind at the time of an utterance, as Chafe defines them, let alone what is expected or unexpected, as Halliday defines them. This is summarized by Prince as follows: “The speaker assumes that the hearer ‘knows,’ assumes, or can infer a particular thing (but is not necessarily thinking about it)” (1981:230).

To illustrate the distinction, a person might start a conversation with, “What do you think of George W. Bush?” and although the addressee would almost certainly be familiar with the current U.S. president, he probably would not be thinking about him at the moment the question is asked. In Haviland and Clark’s terms, *George W. Bush* would represent old information (shared knowledge), but in Chafe’s (and also Halliday’s) terms it would represent new information in the mind of the hearer. While such knowledge would be stored in long-term memory, the speaker could not assume that the knowledge of the president would be active in the short-term thoughts of the hearer, let alone that it would be predictable. The difference, then, between Chafe’s definition and that of Haviland and Clark is largely one of identifiability, to be discussed further in section 3.3.3; however, another difference has to do with propositional vs. referential information.

3.3.2 Propositional vs. Lexical or Referential Information

One reason for the conflicting definitions of the terms *given* and *new* in relation to information structure centers on the definition of information. Lambrecht (1994:47) differentiates between “propositional information” and “elements of information.” The latter he suggests one might also call “lexical information” or perhaps “referential information,” which would include the meanings of words or phrases and the individual referents they express.² Propositional information, on the other hand, is expressed in terms of relationships. Lambrecht claims that “information can strictly speaking only be conveyed relationally, via propositions. Informing a hearer of something means informing him or her of some state of

² I use the term *referents* to refer to discourse entities (i.e. those entities in some real or imaginary world about which people speak), which can be represented in actual speech by nouns phrases, pronouns, or zero anaphora. These manifestations I call *forms of reference* in contrast to the referents themselves.

affairs, i.e. of something which necessarily involves not only participants but also something to participate in" (1994:46). In other words, while the individual parts have meanings, they cannot communicate information in isolation, since they, by themselves, have no power to inform the hearer of a state of affairs. Thus, regardless of whether a particular referent is given or new, its relationship to other parts of its sentence or context makes it part of a larger proposition that conveys some kind of information (however minor) about that referent and the world in which it exists. This propositional information can be given or new.

According to Lambrecht, the confusion arises because it is often the case that certain sentence constituents, particularly subjects, are considered to represent old information, while other parts, usually predicates, are said to introduce new information. He argues that it is not appropriate to segment the information carried by a sentence in that way; rather, the new information that is conveyed is expressed in terms of the relationship between the sentence constituents—new and old (Lambrecht 1994:47-8). He illustrates the problem of segmentation with the simple sentence "She DID it," (emphasis his) in which the forms of each constituent indicate a given status, yet the sentence as a whole represents new information (Lambrecht 1994:48-9). This distinction seems to be partly responsible for the difference between Chafe's definitions of *given* and *new* and those of Haviland and Clark, who, at least in some instances, consider shared knowledge in propositional terms.³ On the

³ Consider the following example, which they break down into given and new propositions: "The jokes Horace tells are awful. Given: Horace tells jokes. New: Those jokes are awful" (Haviland and Clark 1974:513). While the identity of Horace and even the fact that Horace tells jokes are assumed in this example, the most natural, straight-forward reading (in the absence of contextual information to the contrary) indicates that the speaker believes he is communicating a brand-new proposition to the hearer, namely that Horace's jokes are awful. When presenting a new proposition, as in this example, the speaker usually assumes that the listener does not necessarily know, or perhaps has not really considered, the proposition presented, even if the referents expressed by individual constituents are already familiar.

other hand, Chafe and others (including Prince) focus more on individual referents than propositions, and this present study does the same.

3.3.3 Identifiability

The term *definite* has often been used for particular referents that a speaker assumes his audience can positively and exclusively identify. The reason for this is easy to understand, since often these referents are marked in English by the definite article or other determiners that usually convey definiteness. Chafe, however, suggests that *identifiable* would be a better term, because the decision to give the referent definite status goes beyond the speaker's assumption that the hearer knows the referent, but also includes the assumption that the hearer can uniquely identify or "pick out, from all the referents that might be categorized in this way, the one [he has] in mind" (1976:39). Lambrecht (1994) also adopts the term *identifiable*, using it in a similar way.

Identifiability does not include any assumptions about a referent's given-new status as defined by Chafe. The only assumption is that the hearer knows and can uniquely identify the referent, which corresponds closely to Haviland and Clark's "shared knowledge" definition of givenness. While it is true that definiteness (identifiability) tends also to correspond with Chafe's definition of givenness, and indefiniteness (unidentifiability) tends to correspond with newness, this is not always the case. Chafe gives examples of a new referent which is definite and, though rare, a given referent that is indefinite.⁴ Prince (1981)

⁴ In the sentence "I talked with the carpenter yesterday," *the carpenter* is definite, yet the identifiable referent that the NP represents can be new to the discourse (Chafe 1976:43). Examples of unidentifiable yet given referents are much rarer and, according to Chafe, can only be found "in cases where the indefinite referent is different from the referent which established the givenness" as in the example: "I saw an eagle this morning. Sally saw one too" (Chafe 1976:42). In this case the type *eagle* is activated in the mind of the hearer by the first

incorporates the concept of identifiability in her more detailed classification, outlined below in section 3.4, which integrates the above given-new dichotomies.

3.4 Prince's Taxonomy of Assumed Familiarity

3.4.1 Integrating the Given-New Dichotomies

So far we have discussed different binary definitions of given (or old) and new referents, in terms of predictability, saliency, and shared knowledge or identifiability, but nothing has integrated these ideas into a unified system, until now. Prince (1981) incorporates these different and often conflicting ideas into a more detailed, integrated classification system. In doing so, she uses, clarifies, and validates the differing definitions of givenness that were discussed above.

Prince begins by reviewing the three definitions of givenness outlined above—predictability (Halliday), saliency (Chafe), and shared knowledge (Clark and Haviland)—and showing that they are related. She notes that if a referent is predictable, then it must also be in the hearer's consciousness (salient), and if it is in the hearer's consciousness, then the hearer must have some previous knowledge of it or be able to make an inference about it (Prince 1981:231-2). Thus predictability includes saliency, which in turn includes shared knowledge, and so shared knowledge is the most basic or general kind of givenness. Prince, therefore, starts with the notion of shared knowledge, which she renames "assumed familiarity" in order to allow for false assumptions about the hearer's knowledge and to avoid the confusion common with the term *givenness*. From that all-encompassing starting

sentence, so Chafe claims that the second eagle is given even though it is not a particular eagle that the hearer could uniquely identify.

point, she is able to build a taxonomy that includes the conflicting and yet complementary definitions of givenness. To facilitate the integration of these different definitions, Prince, in later work (1992), proposes the terms “hearer-old/new” (shared knowledge criterion) and “discourse-old/new” (salience criterion). Even though these terms were not specifically used in Prince 1981, it will be helpful to keep them in mind and refer to them throughout the discussion.

3.4.2 Evoked Entities

Prince (1981) proposes three basic categories of assumed familiarity: new, inferable, and evoked. The term *evoked* represents given entities that are both hearer-old (given in terms of shared knowledge/assumed familiarity) and discourse-old (given in terms of salience, i.e. already active in the consciousness of the hearer). Most agree that these are given. There are two kinds of evoked entities, those that have been previously mentioned in the text or “textually evoked”—usually shortened to just “evoked,” and those that are evoked by virtue of the extra-textual situation, thus “situationally evoked.” Examples of situationally evoked entities are referents expressed by first or second person pronouns (like *I* or *you*) or the proximal demonstrative *this*, where the identifiability of the referent is tied to the extra-linguistic context of the utterance. Textually evoked referents are often expressed in English by third person pronouns or noun phrases that use the definite article.

3.4.3 New Entities

New entities, on the other hand, are discourse-new, but can be hearer-old or hearer-new. By this breakdown, we can see that Prince has followed Chafe’s terminology by

assigning some hearer-old entities in the *new* category rather than treating them as given. However, she does include Clark and Haviland's terminology in her division of new entities into "brand-new" (hearer-new) and "unused" (hearer-old). Thus, *unused* referents would be those with which the hearer is assumed to be already familiar—hence *given* in Clark and Haviland's terms—but new to the discourse, and therefore, not assumed to be already in the consciousness of the hearer. The classic example of this is the name of someone familiar, such as George W. Bush, as we saw before, but the category is certainly not limited to proper nouns. *Brand-new* referents (hearer-new and thus automatically discourse-new) can be further divided into "brand-new (unanchored)" and "brand-new anchored." Anchored means that the referent is tied to a given entity in the immediate context in which it is introduced. For example, in the phrase "a woman that I work with," *a woman* is brand-new, but is anchored to the speaker (who is situationally evoked, and therefore given) by the relative clause "that I work with." The referent is still brand-new, so the hearer needs to form a new mental representation of it, but the anchor helps to make the mental image more specific, and probably easier to recall. Whether anchored or unanchored, brand-new entities are considered by all to be new entities. They are often but not always expressed by the indefinite article in English.

3.4.4 Old and New Entities in Terms of Hearer-status and Discourse-status

Before we examine Prince's third basic category of assumed familiarity, inferable, it will be helpful to outline the types of discourse entities already mentioned. So far the distinction between hearer-status and discourse-status has produced three types of discourse entities: evoked, unused, and brand-new. As we saw, the status of the first and the last are

not generally disputed, but unused referents can be classified as either given or new, depending on whose terms are used. Prince (1981) follows Chafe in classifying unused entities as *new*—essentially discourse-new. The categories presented above are represented in table 5 below, adapted from Prince 1992:309.

Table 5. Prince's Categories in Terms of Hearer- and Discourse-status

	Discourse-new	Discourse-old
Hearer-new	Brand-new	---
Hearer-old	Unused	Evoked

Table 5 shows nicely how the definitions of Chafe and those of Clark and Haviland line up. Discourse-status represents Chafe's division, and hearer-status represents Clark and Haviland's division. Where discourse-new matches hearer-new or discourse-old matches hearer-old, the authors are in agreement on the status of the referent. But where discourse-new is also hearer-old, they disagree on the status. Prince's use of both criteria demonstrates the validity of both approaches when it comes to classifying referents. Note the empty discourse-old, hearer-new slot. By definition, anything discourse-old (i.e. already introduced in the discourse) also must be assumed to be familiar to the hearer (hearer-old), so discourse-old, hearer-new is an apparent contradiction.

3.4.5 Inferable Entities

Prince's third major category, inferable entities, consists of those that are not evoked, but can be identified via some kind of inference. While evoked referents are preceded in a discourse by a direct antecedent (previous mention), inferable referents are identifiable (and have definite status) via inference from an indirect antecedent. This is illustrated by a

frequently-cited example from an experiment conducted by Haviland and Clark: “We checked the picnic supplies. The beer was warm” (1974:515). The definite article before *beer* indicates that the hearer or reader is expected to be able to identify the beer in question or to be able to infer its identity on the basis of an indirect antecedent in the previous sentence, in this case, *the picnic supplies*. Clark and Haviland call this “building an inferential bridge” (1977:6).

Sometimes the basis for making an inference is stronger, and given elements inherently imply the existence of other elements. These are what Chafe calls “entailments” (1976) or associations with a “schema” (1987). The term *schema* was introduced into psychology by Bartlett (1932), but was previously used by Kant, according to Rumelhart (1980). A schema automatically entails certain elements that are associated with it. The example Chafe gives from the text analyzed in his 1987 article is that of an undergraduate college class. This schema includes: “students, an instructor, teaching assistants, the instructor’s notes, a classroom, [and] a lecture” (1987:30). Another common example is that of a house, which has doors, windows, a kitchen, bathroom, etc., all of which can be inferred from the mention of a house. With a schema one might say that the materials for building an inferential bridge have already been gathered, or perhaps that a shorter bridge will do. Other names that are frequently employed are “frame” (cf. Prince 1981; Brown and Yule 1983; and Jones 1983) and “Stereotypic Assumption” (Prince 1981). Whether the associations are automatic as in a schema or frame or they require more processing on the part of the hearer, the definiteness of all such referents is established by some kind of association or inference.

Prince (1981) proposes two kinds of inferable referents, “containing” and “noncontaining,” the latter being the default, and requiring no further explanation. Containing inferables are a special subset of inferable entities that are inferred from an element that is “properly contained within the Inferrable⁵ NP itself” (Prince 1981:236). Prince gives the example: “one of these eggs,” where the whole noun phrase represents the inferred entity. It is inferred from the given element *these eggs*, which is contained within the larger noun phrase representing the inferable entity. While it is often the case that a member-set or part-whole relationship exists between the inferred referent and the given element, this is not the basis for the classification “containing inferable,” as Prince 1992 makes clear. In reference to containing inferables Prince writes, “the entity which triggers the inference is not, as in the case of [noncontaining] Inferrables, necessarily in the prior discourse, but is rather *within the NP itself*” (1992:307, emphasis hers).⁶

The given-new status of inferable referents is not entirely clear. Consider the beer in our first example. Haviland and Clark would argue that it is given because it is identifiable. However, Chafe (1976) claims that, even though it is identifiable, it is not given. He points out that *the beer* could not be replaced by the pronoun *it* in that context, as it could if the first sentence read, “We got some beer out of the trunk,” which was the case in Haviland and Clark’s direct antecedent pair (1974:514). Prince (1981) leaves inferable entities in a

⁵ There are apparently two acceptable spellings of this word. Prince uses the spelling *inferrable*, and some authors have followed suit. Others, including me, use the spelling *inferable*, except when quoting another.

⁶ Therefore, in the utterance, “Where did you buy these eggs? One is broken,” *one* is an inferable entity but not a containing inferable entity, because *these eggs* is not properly contained within the NP of the inferable referent.

separate category rather than trying to group them with either (discourse-) new elements or evoked (discourse-old) elements. She deals with the subject again in Prince 1992:

Inferables are thus like Hearer-old entities in that they rely on certain assumptions about what the hearer does know, e.g. that buildings typically have doors...and they are like Discourse-old entities in that they rely on there being already in the discourse-model some entity to trigger the inference.... At the same time, Inferables are like Hearer-new (and, therefore, Discourse-new) entities in that the hearer is not expected to already have in his/her head the entity in question (1992:305-6).

In the end Prince concludes that inferables belong in their own category, despite the temptation to preserve binariness. She further notes that what is considered inferable “may itself be subdivided, possibly as a continuum, according to the type and ease of inferencing that is required” (Prince 1992:306). Birner (1994, 1997), however, gives evidence that inferable entities pattern in the same way as evoked entities, particularly in instances of subject-predicate inversion, passivization with “by” phrases, and topicalization.

Prince’s taxonomy of assumed familiarity is represented in figure 2:

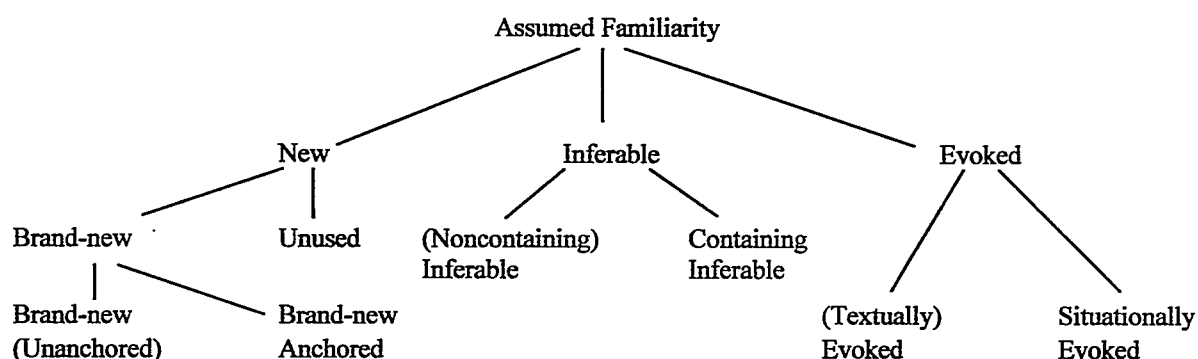


Figure 2. Prince’s taxonomy of assumed familiarity (from Prince 1981:237).

She also proposes a hierarchy which she dubs the “Familiarity Scale” (Prince 1981:245). It is represented as in figure 3, below:

$$\{E/E_S\} > U > I > I_C > BN_A > BN$$

Figure 3. Prince's familiarity scale (adapted from Prince 1981:245): E, (textually) Evoked; E_S, Situationally Evoked; U, Unused; I, (noncontaining) Inferable; I_C, Containing Inferable; BN_A, Brand-new Anchored; BN, Brand-New (unanchored); > means "is more familiar than."

Note that in this case Prince places inferable referents between unused referents and brand-new referents, both of which fall under the discourse-new category of the taxonomy. In support of this scale Prince gives this example (1981:245):

- (35) a. *I bought a Toyota.* (E_S)
 b. *Ellen bought a Toyota.* (U)
 c. *One of the people that work at Penn bought a Toyota.* (I_C)
 d. *A person that works at Penn bought a Toyota.* (BN_A)
 e. *A person bought a Toyota.* (BN)

In sentences b. and c. above it is easy to see that *Ellen* is more familiar (hearer-old) than "one of the people that work at Penn," even though both are apparently discourse-new. *Ellen* is more easily identified, while the referent in c. can be identified by inference only. It should be noted that she presents this scale in terms of familiarity (or identifiability) rather than the overall given-new distinction outlined by her taxonomy, which is more influenced by salience. Notice also the similarity between sentences c. and d.; the semantic difference between the two seems insignificant. Unless "the people that work at Penn" was already given (evoked) instead of just unused, it is difficult to see how c. differs from the brand-new anchored example in d. Perhaps a better example of a containing inferable entity would also be more identifiable and familiar on this scale.

By way of application, Prince notes that if a speaker assumes a certain level of hearer familiarity corresponding to one of the levels on the scale, but chooses instead to violate Grice's Maxims of Quantity by using a form that corresponds to a higher or lower level, the

hearer will be either misled to believe that he does not know the referent when he, in fact, does or confused that he is expected to know a referent that he does not (Prince 1981:245). In chapter 4 we will interact more with Prince's taxonomy and familiarity scale as we apply them to the Adzera texts.

Somewhat similar to Prince's Familiarity Scale is a six-tiered "Givenness Hierarchy" proposed by Gundel, Hedberg, and Zacharski (1993), which they base on referring expressions in several different languages. The major difference is that the latter is an implicational hierarchy in which each level entails all lower levels of givenness, while the levels of Prince's scale are mutually exclusive taxonomic classifications. The focus of Gundel *et al.* (1993) is the degree of givenness represented by different demonstratives, determiners, and pronouns.

3.5 Topic Continuity

The word *topic* has been employed by many people in many different ways. Paragraphs usually have one main topic or theme, even though there may be several subtopics, and whole discourses can have many topics with one overriding, or central, topic or theme. At the sentence level the term usually refers to the basic subject matter of a sentence. Traditionally this has been associated with the grammatical subject of a sentence, but this is not always the case. Sentences can have more than one "topic," though one is usually given prominence over the others via syntactic structure. Givón notes that "topicality" is a somewhat vague, scalar notion, and that topics within a sentence fall into a hierarchy by degree (1983:6). This is particularly noticeable, as Givón points out, in a sentence containing a subject, direct object, and indirect object, that undergoes dative

shifting, so that the indirect object gets promoted to primary object and the direct object gets demoted. Each of the three entities is a topic of the sentence to some degree, and the degree and ranking of each can change with processes like dative shifting. Throughout this paper, unless indicated otherwise, the word *topic* is used not to describe the theme or subject of a sentence, paragraph, or discourse, but rather any discourse entity encoded by a noun phrase or its substitute and found throughout a discourse in any sentence position.

Seeking a way to measure the topicality of discourse entities in an objective way, Givón (1983) developed a quantitative approach to topic continuity. That which is most topical will naturally be most continuous throughout a discourse or portion of a discourse. He notes that there are three types of continuity in discourse: thematic continuity, action continuity, and continuity of topics/participants, the last of which is the most concrete and the one with which we are concerned here (Givón 1983:7).

Our primary interest is the availability, identifiability, or accessibility⁷ of a topic in the consciousness of a hearer at the time of an utterance, or how easily a hearer can identify and retrieve a topic from short-term memory. Givón discusses four factors which affect the degree of difficulty hearers may have in identifying a topic, two of which are particularly relevant to this study: length of absence from the register (or discourse file) and potential interference from other topics (Givón 1983:11). The first can be summarized as: the shorter the gap is between the last previous mention of a topic in a discourse and the current

⁷ Givón uses these three terms interchangeably, and in his usage they are scalar properties. So accessibility here is different from Chafe's (1987) use of *accessible*, which indicates a particular mental status somewhere between active and inactive. In Givón's usage the most accessible topics would be active in Chafe's terms, and the least accessible would be inactive. Likewise, Givón's use of the term *identifiability* corresponds closely with Chafe's usage, except that the latter is a binary distinction rather than a gradient, as it is used here. In this section and in chapter 5, I will use these terms as Givón does.

mention, the easier the topic is to process and identify correctly. The introduction of a new topic is obviously the most difficult to process because a “new file” has to be opened for it. The second factor, potential interference, takes into account the fact that the presence of other semantically similar topics in the immediate context make “correct identification and filing of a topic” more difficult (Givón 1983:11). These two factors form the basis for two measurements of accessibility (or identifiability), to be discussed below.

Givón seeks to measure not just topic accessibility, but topic continuity in general. He lists several assumptions that relate the more specific idea of availability to the more general notion of continuity (Givón 1983:12):

- a. “What is continuing is more predictable”
- b. “What is predictable is easier to process”
or conversely
- c. “What is discontinuous or disruptive is less predictable”
- d. “What is less predictable, hence surprising, is harder to process.”

Givón (1983) proposes three separate quantitative measurements of topic continuity. Two relate specifically to accessibility; they are referential distance and potential interference. The third measurement, topic persistence, deals with how long a topic remains active in the discourse. This study employs all three measurements, which are briefly outlined below.

Referential distance (RD), also called “look-back,” is defined as the gap between the last previous occurrence of a referent and the current occurrence in question, and is expressed as the number of clauses to the left. It is a measure of topic accessibility. Low values indicate accessible, or easily identifiable, topics, while high values indicate inaccessible topics.

Potential interference (PI), also called “ambiguity,” measures the potential that other referents in the immediate context have to create ambiguity and thus make positive identification of the referent in question difficult. The question that this measurement seeks to answer is, “If the minimum possible markings were used, could the referent be confused with another referent with similar semantic attributes in the immediate context?” This measurement works closely with RD to assess availability.

Topic persistence (TP), sometimes referred to negatively as “decay,” is a measure of the longevity and thus importance of the topic. Important topics tend to remain as topics longer in subsequent discourse. Givón has employed two different measurement methods for topic persistence, to be discussed later.

We will discuss this approach, its measurements, methodology, and predictions in more detail when I apply it to the Adzera texts in chapter 5.

3.6 Participant Reference

Participant reference, as the name implies, is concerned with the forms of reference applied to primarily animate participants, a subset of topics in general. This approach ignores the props in a discourse and focuses the researcher’s attention on the participants only. In some texts inanimate objects can play a vital role such that they function more like participants than props (Currier 1977:21-3); in such cases, they should be included. Most of the time, however, the participants are primarily if not exclusively animate beings. On the other hand, there are cases in which an animate referent (even a human referent) functions as a prop rather than a participant by virtue of the fact that the referent plays a minor and completely passive role (Currier 1977:21-3).

Many have written from various perspectives on the subject of participant reference or participant tracking, including, but certainly not limited to Dooley and Levinsohn (2001), J. Grimes (1975), Levinsohn (1994), and Longacre (1983, 1989, 1995). Levinsohn (1994; cf. Dooley and Levinsohn 2001) proposes a methodology for analyzing participant reference. Although the present study is concerned primarily with the reference forms of animate participants in Adzera oral narrative, it examines inanimate entities as well, and follows the quantitative approach proposed by Givón above. The animate participant and inanimate prop distinction is one that we should keep in mind in chapters 4 and 5.

3.7 Related Issues

The above overview by no means covers, nor even begins to cover, the vast amount of literature and ideas written about the related areas of information structure, topic continuity, and participant reference. Many authors in their discussion of given and new information—the assumptions made by speakers regarding their listeners' states of mind—write about commonly employed strategies or structures called “presupposition and assertion/focus” and “topic and comment,” for example. Another common subject of study is the phonological marking of focus, contrast, etc., via stress and intonation. Linguists commonly refer to “intonation units” (Chafe 1987) or “tone groups” (Halliday 1967). All of these concepts are beyond the scope of this study as each would certainly comprise a detailed study on its own; therefore, they have been omitted from the above discussion and will not feature in the rest of this paper.

CHAPTER 4

ASSUMED FAMILIARITY OF DISCOURSE REFERENTS

4.1 Introduction

The last chapter gave an overview of the literature that has been written about the given-new status of information and referents within a discourse. This chapter will apply the terms and methodology of one contributor in this field—Prince. We will examine two Adzera texts, classifying the referents according to Prince's (1981) taxonomy of assumed familiarity, and showing the distribution of each kind of referent in subject and non-subject positions. First, however, we will discuss potential problems of assumed familiarity as illustrated by two examples in one of the Adzera texts.

4.2 Potential Problems of Assumed Familiarity

Human beings have the ability to communicate amazingly well, even when topics of conversation or particular referents are new to the hearer. Languages employ various methods—articles, demonstratives, pronouns, descriptive phrases—to mark a referent as new or given, thus helping the hearer to assign the proper identity to the referent. Most of the time speakers make good use of the rules and features of their languages, enabling them to communicate well. Sometimes, however, a speaker refers to a participant or topic as though already active in the mind of the hearer, but the referent is new to the hearer (at least in the hearer's present consciousness). The hearer has trouble positively identifying the referent,

and miscommunication results. In such situations it is evident that the speaker's assumptions about the hearer's state of mind were not correct.

The reader will remember that Prince suggests that the term *assumed familiarity* is more appropriate than *shared knowledge* because what the speaker assumes is shared knowledge is not necessarily so (1981:232-3). That the speaker may have false assumptions about the hearer's familiarity with a participant or topic is easy to see in every day life. How many times in normal conversation does a listener need to interrupt the speaker to ask for clarification on something that was said or left unsaid? It is human nature to forget that what one has in mind is not necessarily what is in the present consciousness of the hearer.

This is particularly true in the impromptu telling of oral stories. When telling a story, the storyteller has not rehearsed the words in his mind. Rather he is perhaps remembering events as they occurred and retelling them on the fly. It is easy for him to forget to give proper introductions, and to remember that what he has in mind is not necessarily on the minds of his audience at that moment. Admittedly, if the story is being recorded, he may be more careful than he would be in a more casual setting, but on the other hand, he would most likely be more nervous and apt to make mistakes as well. It is just such an environment that will cause discrepancies in assumed knowledge, leading to unclear references.

When the audience is large and diverse, the storyteller's task becomes more complicated. What assumptions will he make regarding the knowledge of the hearers? There may be varying amounts of shared knowledge between the speaker and different members of the audience. Perhaps there are people of different age groups or generations, people from different villages, people from different cultures, people with different levels of

education, or other distinctions that would divide people into insiders and outsiders where a particular item of assumed knowledge is concerned. Now not only will the speaker need to be careful to activate unused referents, but he will most likely need to introduce referents that are brand new to some. To whom does the speaker tailor his words in such cases? Most speakers would presumably accommodate the outsiders as much as possible with extra explanations when needed, but they would also be inclined to leave out certain details with which most of the hearers are already quite familiar. We saw that Grice's Maxim of Quantity speaks directly to this issue, suggesting that to be cooperative a speaker must make his contribution informative enough but not more informative than necessary. Thus, the speaker faces a dilemma. In some cases, he may choose to tailor his story for those who are "in the know," and let the others inquire if they do not understand. It is likely, however, that in many cases of omitted information the speaker simply forgets that certain members of his audience may not share knowledge of a particular referent that is very familiar to both him and the majority of his hearers.

I am an outsider to Adzera culture, and in spite of all the ways that my Adzera friends have accepted me and regardless of how long I live and work among them and how much I adapt, I will never obtain the full cultural knowledge shared by those who were born and raised in that culture. Nor do I have the historical and geographical background knowledge about their area that they possess. Most Adzera speakers assume that my knowledge is limited, and they try to be accommodating. However, at times people may assume that I know more than I do. But even among the Adzera people themselves there are different dialects, district groups, villages, and clans spread over the broad valley. What is familiar to

members of one group may not be to members of another. This increases the likelihood of a discrepancy between the assumed familiarity of a hearer with a particular topic and the hearer's actual familiarity.

4.3 Apparent Problems of Assumed Familiarity in an Adzera Text

Consider the following example taken from text A. The storyteller uses the third person plural pronoun ('they') in a context where its intended referent is not clear. Several clauses later he employs a descriptive noun phrase in a construction that appears to equate this new referent with the previous 'they'. Does 'they' refer to a group previously mentioned, as it normally would, or does it refer to the group made explicit several clauses later? It is possible that the speaker had this latter group in mind, but did not make it explicit initially. Note how the construction in sentence A-76 can affect the interpretation of 'they' in sentence A-72.

(36) A-71 'We went to Gimigim, a village of the Wagiasa people.'

A-72 *Aga mpa arigi, orait, ribigi ya?... irurub rib Saum.*
 1pl(ex) stay that.one alright 3pl (um) R-baptize 3pl Saum
 'While we were living there, they baptized the people of Saum.'

A-73 '(They) baptized everyone.'

A-74 '(They) baptized the people of Saum, Sukurum, Fayang, Ngariawang, Sirasira, who else...Papuai, Sisuk, Mugaba.'

A-75 'All of them (they) baptized at Saum village.'

A-76 *Ramarj' ruas ani ifan da irurub.*
 father-1pss PLUR here R-go and R-baptize
 'Our fathers here went and Ø baptized Ø.'

In English, sentence A-76 could have only one interpretation, since the only time the subject of a clause may be ellipted (i.e. omitted) is in the second subject position of

coordinate clauses¹ within the same sentence, and only if it is co-referential with the first subject. Therefore, in English the subject of ‘baptized’ must be ‘our fathers here’. This would be the most natural interpretation in Adzera as well, and, divorced from context, the only interpretation possible. This interpretation, which I shall designate Interpretation #1, is represented as follows:

- (37) A-76 [*Ramaŋʹ ruas ani*]_i *ifan da* Ø_i *irurub* Ø_j.
 father-1pss PLUR here R-go and R-baptize
 ‘[Our fathers here (the church elders)]_i went and Ø_i baptized (them)_j.’

Following Interpretation #1 requires that the ‘they’ in A-72 is co-referential with, or at least inclusive of, *ramaŋʹ ruas ani* ‘our fathers here’ in A-76. If such is the case, then the speaker has utilized the third person pronoun (‘they’)—normally reserved for given referents—without prior reference to the participants. We must then assume that the speaker initially failed or chose not to make clear the referent who was doing the baptizing (in A-72), but later corrected his “error” by making the referent explicit (in A-76).

However, according to an Adzera language helper who was unfamiliar with the story, this sentence could have two other possible interpretations besides Interpretation #1. One is that ‘our fathers here’ (i.e. the local church elders) went to witness the baptism, but another group did the baptizing. We can represent Interpretation #2 in the manner shown in (38):

- (38) A-76 [*Ramaŋʹ ruas ani*]_i *ifan da* Ø_k *irurub* Ø_j.
 father-1pss PLUR here R-go and R-baptize
 ‘[Our fathers here (the church elders)]_i went and (they)_k baptized (them)_j.’

¹ In the case of a string of coordinate clauses, ellipsis may occur in the subject position of any or all clauses except the first.

Under Interpretation #2, *ramanj' ruas ani* 'our fathers here' would not be co-referential with either the subject or the object of the second clause. While such an interpretation is possible in the larger context, it seems unlikely because of the high potential for ambiguity. Barring other factors that would require a different interpretation, a null subject should be interpreted as co-referential with the subject of the preceding clause, if semantically compatible (Burquest 2001:275).² Only in cases where the baptizers were clearly identified beforehand could this sentence, as indexed in (38), be used without ambiguity. On the other hand, if the speaker thought he had made it clear who did the baptizing, then he could possibly use such a sentence with the intended meaning of Interpretation #2. The advantage of this interpretation is that 'they' in sentence A-72 would refer to a group previously mentioned—most logically the people of Wagiasa (based purely on the most recent mention in the discourse). The use of the pronoun would then fit with the status of the referent.

The third possible interpretation, according to the language helper, is that people from the speaker's father's generation went up from his village and were baptized. In that case *ramanj' ruas ani* 'our fathers here' would not be referring to the church elders of that time, but instead would refer to people of that generation who would have been new converts at the time. With Interpretation #3 the subject of the first clause ('our fathers here') is also the object of the second clause, as example (39) shows:

² Consider sentences B-22 through B-25 (appendix B) and the semantic factors that guide the interpretation of ellipted participants. In the physical interaction between two highly salient participants, the human referent is consistently ellipted (even in places where it seems confusing to do so), while the non-human animate referent is referenced with an overt noun (except its ellipsis as DO in B-22). It is actually the semantic properties of the verbs in B-23 and B-24 that alert the reader that the ellipted references are to the human participant. See also examples of ellipsis in turn-taking within the dialogue in B-32 through B-35.

- (39) A-76 [*Ramaŋ'* *ruas ani*]_i *ifan da* Ø_k *irurub* Ø_i
 father-1pss PLUR here R-go and R-baptize
 '[Our fathers here]_i went and (they)_k baptized (them)_i.'

This interpretation shares the same advantage of the previous one, namely that 'they' would refer to a previously mentioned referent, making it consistent with normal pronoun usage. It is also more syntactically plausible than Interpretation #2 because the ellipped object of 'baptized' is co-referential with the last explicit referent. Here it would be helpful to compare it to the previous sentence, which contains a preposed direct object, as we saw in section 2.3:

- (40) A-75 [*Santan aŋu*]_i *da* Ø_j *irurub* *t_i imiŋ'* *rib Saum gampan*.
 all only and R-baptize R-at 3pl Saum village-3pss
 '[All of them]_i they_j baptized at Saum village.'

- A-76 [*Ramaŋ'* *ruas ani*]_i *ifan da* Ø_{j?} *irurub* Ø_{i?}.
 father-1pss PLUR here R-go and R-baptize
 '[Our fathers here]_i went and (they)_{j?} baptized (them)_{i?}.'

In sentence A-75, *t_i* marks the position where the object would normally be found. The conjunction *da* 'and' sets the noun phrase *santan aŋu* 'all of them' apart as a preposed object rather than the subject of *irurub* 'baptized'. In sentence A-76 *da* 'and' serves a different purpose, namely to join coordinate clauses. If the verb *ifan* 'went' were not present in the first clause of sentence A-76, then it would be identical in structure to A-75 and *ramaŋ'* *ruas ani* 'our fathers here' would be most naturally interpreted as a preposed object. While it is still possible that it is co-referenced with the ellipped object of 'baptized', which the similar structures might possibly support, one cannot deny that *ramaŋ'* *ruas ani* is more naturally co-referenced with the ellipped subject. This is especially apparent when we consider that the

speaker could have, and probably would have, eliminated the ambiguity by making the subject of *irurub* 'baptized' explicit if he had intended either of the latter two interpretations.

When we have investigated all the textual clues, we can turn to the historical context of a true story such as this one. This, of course, assumes that enough is known about the historical situation in which the events took place. In this case the background information supports Interpretation #1. There was a mission station not far from the storyteller's village, and evangelists from the area went out to the more remote areas to work, including the storyteller's father (a fact which we can actually glean from the text itself). The mountainous regions to the north and south of the broad plain where the narrator lives would have been the more remote areas targeted for evangelism. This fact is supported by the text, since the main character's target villages were in one of these mountainous regions. These observations alone make *raman' ruas ani* 'our fathers here' the probable baptizers and equally unlikely co-recipients of a baptism taking place in the mountains.

Interpretation #1 is clearly the best choice for syntactic and contextual reasons. If the syntax of the sentence is examined in isolation, it is the only interpretation possible. The lack of any other positively explicit referent in the previous several sentences and other contextual clues, both in and out of the text, support this interpretation. Furthermore, as we shall see in chapter 5, the introduction of the noun phrase *raman' ruas ani* 'our fathers here' would normally require a more explicit subject for *irurub* 'baptized' if the intended referent was not 'our fathers here' (but cf. footnote 2, above).

Even though information structure theory suggests that the third person plural pronoun ('they') in A-72 should have a noun phrase antecedent, there is enough evidence to

show that such is not the case in this instance. In focus is the object of the action, those who received the baptism. The speaker either inadvertently failed to make the agent explicit initially because he was focusing on those who were baptized, or perhaps more likely, he chose not to specify the agent in order to focus on those who were baptized, because Adzera does not have a passive. The constructions in A-72 through A-75 could then be functioning like passives.³ Staley (1995) describes instances of third person verb affixes and/or pronouns introducing new referents in Olo, a language of Papua New Guinea that also lacks a passive construction. Like the passive in many languages, Staley sees this usage in Olo as a grammatical device for demoting these agents and keeping the conscious awareness of them to a minimum (1995:183). If the speaker in our example is employing this kind of strategy, then the identity of the third person agent in these sentences remains unspecified because it is not as relevant to the story as who was baptized.⁴ A non-specific⁵ reference in A-72ff would still allow any one of the three interpretations of A-76, but for reasons presented above, I believe Interpretation #1 is still the best one. Therefore, I consider *ramarj' ruas ani* 'our fathers here' to be the subject of *irurub* 'baptized', and, therefore, co-referential with the previous 'they'.

³ Sentence A-74 (see appendix A) introduces the objects of 'baptize' with the object marker *i*, though it is not necessary. This looks similar to the emphasis sometimes used with locative objects, which S. Holzkecht describes (1986:134), and may further support a passive-like interpretation.

⁴ Responses received from language helpers indicate a vague reference in A-72. The fact that three different interpretations of A-76 are possible demonstrates the vagueness of the preceding reference. Furthermore, the language helper who gave those interpretations of A-76 said in reference to A-72 that baptism 'is the work of missionaries and church elders only'. The other language helper said that 'they' refers to "those who baptize people," not specifying a particular group within the text, and that 'our fathers here' refers to "Sangan church elders," not necessarily connecting the two.

⁵ It is "non-specific" in the sense of unspecified but not non-referential, because it does refer to a particular group, even though the identity of that group may not be relevant to the story.

There is another example in the same text of a vague reference involving 'they'.

Consider the sentences in (41):

(41) A-18 'He (*ramanʔ*) went and stayed with Ngawai in Sukurum.'

A-19 *Raiyi da waʔ ribigi ini nan*
later TIME (um) 3pl-there R-say talk

da itip iyu ramanʔ iba fawaʔ Saum.
and R-again R-take father-1pss R-come go-arrive Saum

'Later, they spoke and brought my father back to Saum.'

Syntactically, one would expect 'they' in A-19 to be co-referenced with the subject of A-18, but in this case that is *ramanʔ* 'my father', which is not only a singular referent but is also the object of the second clause of A-19. This interpretation could only work if he had others (perhaps including Ngawai) acting with him in making a group decision that directly affected him,⁶ but one still does not know who those others are, so the identity of the group remains unclear. The pronoun 'they' could refer to the Sukurum people (as one language helper suggested), to the Saum people, to both, or to some other group that is not specified. Another language helper suggests that it refers to the group who originally sent the speaker's father to do evangelism, but this group is not mentioned anywhere in the text, nor is there mention of his being sent, even though it is undoubtedly the case that a congregation or church group sent him.

The point is that the reference is quite vague, making it impossible to pin down a clear antecedent with any certainty. In an attempt to do just that I could analyze it as the Sukurum people, inferred from the previous sentence (A-18). But the pronoun points no more clearly to that group than it does to any other entities previously mentioned or

⁶ Decisions in Adzera culture are made via group consensus by the married men (cf. section 1.4).

unmentioned. Here the reference could be to any, all, or none of the above, but again the specific identity is not relevant to the storyline. Therefore, the speaker is not violating Grice's maxims by not providing enough information; rather, he is applying them by not providing more information than is required or giving irrelevant information. He assumes that the hearer can infer any relevant information regarding the identity of the agent (which may be no more than "the people responsible for those decisions"), even if the referent cannot be exclusively identified. Indeed, when I heard, transcribed, glossed, and re-read the story I had no trouble at all with the reference until I tried to analyze its given-new status and pinpoint an antecedent. The main point is that the narrator's father was sent to Saum, regardless of who sent him. In the other example, the main point is that those groups of people were baptized; who did it does not really matter, though the author apparently decided afterwards to specify this.

The analyst then faces the dilemma of how to treat these vague references. Does one count them as referents or dismiss them as non-referential? They are technically not non-referential, because they certainly denote specific groups that made specific decisions or performed specific actions at specific points in time, whether or not their identities were actually specified by the narrator. However, the speaker's use of the pronouns is more attributive than referential (cf. Donnellan 1966), so while the pronouns "denote" specific entities, they do not specify or "refer to" those specific entities, in Donnellan's terms.⁷ Therefore, these references fall between referential and non-referential, if that is possible. While either inclusion or exclusion of them would

⁷ Donnellan (1966) applies these terms to full noun phrases, not pronouns. However, the terms also seem to apply to the vague use of pronouns in which the speaker does not commit himself to a particular identification.

probably be valid, I include them as referents, because they do denote specific entities, and I believe the one reference is subsequently specified.

In these cases, then, the source of the problem is apparently not the speaker's assumptions about the hearer's state of mind, as it often is in unrehearsed oral narratives. Rather, if indeed there is a communication problem, it may be because the hearer (or maybe just the analyst) assumes that the speaker expects him to be able to uniquely identify these referents, when, in fact, this is not the case. In these examples, it is likely that the speaker uses vague agent references to downplay the agents and highlight the predicates in the absence of a passive construction.

4.4 A Classification of Referents in the Adzera Texts

4.4.1 Methodology

In this section I classify all the referents in my texts according to Prince's taxonomy of assumed familiarity. The method employed here examines only the overt forms of reference—full noun phrases, proper nouns, pronouns, demonstratives—that represent discourse referents. The first step was to list each overt reference in the text, making two lists: animate referents and inanimate referents. References embedded within other referring expressions⁸ that represent different entities were not listed separately nor counted. For example, in the noun phrase *Augustin bajin* 'Augustin's hand', Augustin is a given, animate referent, and his hand is an inanimate referent that is

⁸ In this thesis the term *referring expression* has the general definition employed in pragmatic studies, namely any overt reference to a discourse entity. Principles and Parameters theory uses a narrower definition that distinguishes referring expressions—i.e. full noun phrases and proper nouns—from pronominal references. In the above case, both definitions would apply.

inferable via human anatomy; in this case the word *Augustin* was not listed again as another reference to Augustin, but was considered to be only part of the larger referring expression. The second step was to designate each overt reference as a grammatical subject or non-subject, and classify its referent (at the time of the utterance) according to Prince's (1981) taxonomy of assumed familiarity. I assigned each manifestation of a referent one of the following classifications: Brand-new (BN), Brand-new Anchored (BN_A), Unused (U), Inferable (I), Containing Inferable (I_C), Evoked (E), and Situationally Evoked (E_S), as discussed in chapter 3. The final step was to count, for both animate and inanimate referents, the number of each of the above seven classes occurring as subjects and non-subjects.

This methodology does not account for the most salient of Adzera referents, those which are ellipted. If ellipted referents were counted as evoked, then evoked entities would comprise an even greater percentage of the total referents. I will account for ellipsis in the next chapter. In addition this analysis excludes clearly non-referential entities, for example, those that represent a type of animal in general such as eagles (A-33), but do not represent a specific instance of an eagle or a group of eagles. Finally, I exclude references to propositional information instead of discourse entities, even though Prince (1981) included some expressions that represent propositions.

I will use the portion of text in example (42), below, to illustrate the above methodology, and will reexamine the same portion in chapter 5 to highlight the differences between the two approaches.

- (42) A-48 *Da Mugusa₁ irunt iba.*
 and Mugusa R-run R-come
 'Then Mugusa came running.'
- A-49 [*Gurump narungan bijangan*]₂ *Mugusa₃.*
 Gurump child-3pss name-3pss Mugusa
 'Gurump's son's name was Mugusa.'
- A-50 *Da isalim arigi₄, ini bina₅,*
 and R-send that.one R-say like.this
 '(He) sent him, saying,'
- A-51 <<*Warunt da wani da mision₅ i ya₆*
 IMP-run and IMP-say to missionary Comp (uh)
Sijkom₆ buy₇ a₈... nabia₉ nais a₁₀, mision₇
 Sinkom (FUT) (IR-) FUT-come FUT-kill (uh) missionary
da naga da gubu₁₁ arani.
 and FUT-eat TIME day this.one
 '<<Run and tell the missionary that Sinkom will come and kill him
 and eat (him) today.'
- A-52 <<*Da natip rungan₈ ma nafan*
 and.so HORT-prepare self-3pss or HORT-go
ya nagup.>>
 and HORT-hide
 '<<So (he) should get himself ready or go and hide.>>'
- A-53 *Da ramay₉ ini da dzi₁₀,*
 and father-1poss R-say to 1sg
 'And my father said to me, ...'

I counted ten overt references in this portion of text, excluding the time expression *gubu₁₁ arani* 'today' in A-51. I have marked them with numerical subscripts, and I consider numbers 1, 2, 6, and 9 to be subjects, and the rest to be non-subjects. The first reference, in sentence A-48, is the first mention of the referent Mugusa. Looking at A-48 alone, the

referent appears to be unused, since the narrator uses only a proper noun. However, the description in A-49 reveals that the narrator does not expect the hearer to know the referent in the previous clause, so I classified *Mugusa* in A-48 as brand-new. The second reference, in A-49, is the NP *Gurump narunggan bijangan* ‘Gurump’s son’s name’. This is an example of multiple embedding. The reference to the character *Gurump* is evoked, but this approach does not evaluate *Gurump* or *Gurump narunggan* ‘Gurump’s son’ as referents here because they are part of the larger referring expression. The whole NP here in A-49 is more difficult to analyze than the previous example. Since the NP is equated with the name *Mugusa*, which was already mentioned, one could call it an evoked entity with an additional attribute, namely the specific identity of the previous referent based on his relationship to *Gurump*, a given referent. But the relationship information provided is new and necessary for proper identification of the referent. Others might classify it as brand-new anchored because the specific identity associated with the name is new and “anchored” by the given character *Gurump*. This view might work to capture the narrator’s intentions and assumptions if we treat the sentence as a fresh start and replacement for the previous sentence. I classify this as containing inferable referent, because it is not necessary for the author to first assert that *Gurump* has a son or that his son has a name; rather the NP ‘Gurump’s son’s name’ assumes this information, even though little inference is necessary. The second half of the equation in A-49 is the name *Mugusa*, and I consider it to be an evoked entity. The fourth reference, *arigi* ‘that one/him’, is found in A-50, and is also evoked, referring back to *Mugusa*.

The next four references are found inside the embedded speech in A-51 and A-52. In A-51, numbers 5 and 7 consist of the unmodified noun *mision* ‘missionary’, which refers to

the main character, and is, therefore, evoked. Number 6 is the proper noun *Sinjkom*, which is already evoked from previous mention. The eighth reference (in A-52) is the reflexive pronoun *rungan* ‘himself’, and is evoked.

The last two references are found in A-53: *raman* ‘my father’ and *dzi* ‘me’. Both are evoked by virtue of numerous previous mentions, and the latter can be said to be situationally evoked, as well.

4.4.2 Specific Examples of Referent Classes

In the above portion of text, I have illustrated what kinds of entities are counted in this chapter, and have classified the ten referents I found there. In this section I will present each referent class separately and give specific examples of each from the Adzera texts.

4.4.2.1 Examples of Brand-new Referents

As was noted in chapter 3 and as the name suggests, brand-new entities are those of which the speaker assumes no prior knowledge whatsoever on the part of the hearer. It does not mean that the hearer could not have had prior knowledge, just that the speaker does not assume such knowledge and so presents the referent in such a way that someone totally unfamiliar with that particular referent can still understand the reference. This is evident in example (43):

- (43) A-29 *da Sinjkom, garam tsira' manan birangan, isu ruruai*
 and Sinkom man big a name-3pss R-become gov't.officer
 ‘and Sinkom—(that’s) the name of one of the leaders—became
 the local officer for the colonial government’

In example (43), a proper noun is followed by a parenthetical description, indicating that the speaker does not expect his audience to positively identify the character by name alone. The speaker apparently started to refer to the referent by name only, but then caught himself and added the parenthetical description. If a parenthetical description like the one in (43) can be a valid anchor, then I would call this brand-new anchored. We saw above the first mention of the name *Mugusa*, which was brand-new, even though it initially appeared to be unused. This occurs at the peak of the story, and in the emotion and excitement the storyteller leaves out important details and has to repair the problem as he goes. This is not the case in the following example:

- (44) B-4 *Ifan da itsaja mpu maran manan da irunt in iru?*
 R-go and R-see water spring a and R-run 3obj R-down
 'He went along and saw a spring of water and followed it
 downstream.'

In example (44), the hearer, though familiar with the type *spring*, would need to create a new "file" in his mind for this specific example, so it would be hearer-new. Note that the equivalent of an indefinite article is found in both examples (43) and (44); however, there are many examples in which this is not the case. Brand-new referents in these texts take the form of full NPs with modifiers (with or without indefinite determiners), proper nouns (with an appositional description in (43)), or unmodified common nouns.

4.4.2.2 Examples of Brand-new Anchored Referents

The reader will recall that brand-new anchored referents are a special type of brand-new referent which is tied to a given referent by virtue of a given element within its referring expression. Brand-new anchored referents are always full NPs, as they require an internal

anchor. In example (44), above, *mpu maran marjan* ‘a spring of water’ was introduced; in example (45), below, it appears again in what could be either a complement clause or a relative clause (cf. section 2.5.4). If we consider the construction to be the latter, then the above referent serves to anchor a brand-new referent.

- (45) B-5 *Ifan da itsarja fugai marjan*
 R-go and R-see crocodile a
[mpada mpui maran arigi wasa' gan].
 sit-NMZ water spring that middle 3pss
 ‘He continued on and saw a crocodile [sitting in the middle of that spring-fed stream].’

Whether it is actually a relative clause remains to be seen, and should be pursued in further study. For illustrative purposes I will consider it to be a relative clause; otherwise, it would not be a good example of a BN_A referent, as the new NP would simply be *fugai marjan* ‘a crocodile’. Treating it as BN_A rather than BN does not significantly affect this analysis.

Often speakers use relationships to themselves or others, including kinship terms, to anchor brand-new referents, as in (46):

- (46) B-2 *Dzi ini fisa i nan dzi bunga' Maramai*
 1sg R-want tell-NMZ Obj. story 1sg inlaw-1pss Maramai
[mpada miŋ'a wa' Riara].
 stay-NMZ at-NMZ (um) Riara
 ‘I want to tell you a story about my father-in-law Maramai
[who lived in Riara].’

Here the narrator is starting a new discourse and introducing the main character. The description *dzi bunga'* ‘my father-in-law’ anchors the referent to something familiar—the speaker himself—via the relationship between them. The relative clause ‘who lived in Riara’ adds additional information about the referent, and also serves to link the new referent to the

hearer-old entity *Riara*, even though both entities are discourse new. Here is another example of a kinship relationship as an anchor.

- (47) B-43 *Da numuntugan Kakarak da wa? Ati? itsaja...*
 and bro.in.law-3pss Kakarak and (um) Ati R-see
 ‘And his brothers-in-law Kakarak and Ati saw it...’

Incidentally, four out of the five brand-new anchored referents in the two Adzera texts are anchored via kinship relations, the exception being the first example given here, which remains questionable.

Some might argue that the NPs “anchored” by kinship relationships are containing inferables instead, similar to ‘Gurump’s son’s name’ in A-49 of (42). One could not infer the fact that Gurump has a son any more than the fact that Maramai has brothers-in-law, and yet the former appears to assume slightly more than the latter, based on the ways they are presented. On the other hand, one can infer that a person has a father, so I consider the first mention of *ramaj* ‘my father’ to be inferable. With the others, one can infer that a person might have a brother, a son, or in-laws. So these referents fall on a continuum. Should all of them be treated as containing inferable, or is it valid to call some of them brand-new anchored? If some are BN_A, which ones? My treatment of these referents as BN_A is tentative, and I am open to calling them I_C. The classification of referents like these demonstrates the subjective nature of this kind of analysis; much of it is based on intuitions, which can change with a fresh look at a later point in time. It also demonstrates the similarities and potential for confusion between containing inferable and brand-new anchored referents. Both contain a new element and a given element that serves as a link. Recall the confusing similarity between Prince’s examples of containing inferable and brand-new

anchored referents given in example (35) in chapter 3 (from Prince 1981:245): *One of the people that work at Penn bought a Toyota.* (I_C) vs. *A person that works at Penn bought a Toyota.* (BN_A). The first presupposes that people work at Penn. Does the second also presuppose this? It does not directly, and yet the same assumption seems to hold.

4.4.2.3 Examples of Unused Referents

Unlike brand-new entities, unused referents are already known to the hearer, but have not yet been mentioned in the discourse. Proper nouns commonly fall into this category as in example (48) below:

- (48) A-13 *Ini fada Dzigintsuan...*
 R-want go-NMZ Dzigintsuan
 ‘He wanted to go to Dzigintsuan...’

However, unused referents are not limited to proper nouns:

- (49) B-3 *Da inaj gum Fraide sib...*
 and R-do work Friday already
 ‘He had finished the Friday community work...’

Everyone in Adzera culture would be familiar with the custom of setting aside certain Fridays for community work. Another example is the word *ruruai* ‘local officer for the colonial government’ (from *Tok Pisin: luluai*) in example (43) above. All of these entities are assumed to be familiar to the hearer but new to the discourse. Hence Prince’s designation “unused.” In addition to proper nouns, which are the most common type, unused referents can appear as modified or unmodified common nouns.

4.4.2.4 Examples of Inferable Referents

In this section we will look at non-containing inferables, and examine containing inferables in the next. Text B has some good examples of inferable entities as illustrated below:

- (50) B-7 *Iwarjun sib da ini apij'a narun*
 R-pregnant already and R-want give.birth-NMZ baby-3pss
da iyab ifan ifari git
 and R-go.up R-go R-build nest
 'She was pregnant and wanted to give birth to her babies
 [i.e. lay her eggs], so she went up and built a nest.'

In this sentence, *narun* 'her babies/offspring' is inferred from the fact that she is pregnant, and *git* 'nest' is inferred from general knowledge about crocodile birthing. One might say that the crocodile-birthing frame has been activated. The first reference in (50), *narun* 'her babies/offspring', might belong in the I_C category because of its possessor agreement marking; I have included those with just agreement to be non-containing, and those with more explicit possessors to be containing. The next example is inferred from an Adzera culture hunting frame.

- (51) B-20 *Da araja ba munti da ifinti sinan*
 and 3sg come stand and R-charm spear
 'He came and stood and put a ritual blessing on the spear.'

There is a similar example in A-7, in which *taramapruf* 'metal-tipped spear' is inferred from a character's attempt to spear a pig, and from the verb *intarj* 'speared'.

I also treated the vague uses of 'they' discussed in section 4.3 as inferable entities, because any information that could not be inferred (from cultural knowledge or textual context) was not essential to the hearer's understanding of the story. The crucial information

in one case was that the main character was sent to work in a particular village, regardless of who sent him. We do know from the pronoun that a group sent him, and we can infer that that group was normally responsible for that kind of decision. With a bit of Adzera cultural knowledge, we can further infer that the decision was made by church and community leaders probably reaching a decision by consensus, but this is not important to our understanding of the story, so the speaker does not provide any more detail. In the other case, the main point is that certain groups were baptized, and we can infer from cultural knowledge that the baptizers must be missionaries or church leaders, which we find to be the case several clauses afterwards (if we accept Interpretation #1). Even though we can infer these things, probably the only crucial information was that someone performed those actions. So, in addition to modified and unmodified common nouns, inferable entities can be pronouns in some cases.

4.4.2.5 Examples of Containing Inferable Referents

The texts also contain examples of containing inferable referents—those whose referring expressions contain a reference to the given entities by which they are inferred. The narrator of text A introduced the place name *Dzigintsuan* as an unused referent (example (48) above); consider now the following example which is inferred via the givenness of *Dzigintsuan*:

- (52) A-13 *da ifa sau garam Dzigintsuan gampan.*
 and R-go look.for man *Dzigintsuan* village-3pss
 ‘So he went and looked for the village(s) of the Dzigintsuan people.’

It should be noted that in Adzera culture (and, I suspect, in most of Papua New Guinea), place names and people names are inextricably linked. The name *Dzigintsuan* can refer to a

group of people or to the place where they live. People usually use such place names for one main village and all of the surrounding hamlets associated with that village and the people who live there. Thus *garam Dzigintsuan* ‘the people of Dzigintsuan/the Dzigintsuan people’ should at least be classified as containing inferable from the first mention of *Dzigintsuan*, if not simply evoked⁹. From *Dzigintsuan* to *garam Dzigintsuan gampan* ‘the village(s) of the Dzigintsuan people’ is, in my opinion, more than just a synonymous restatement, as the former includes the people as well. The latter reference is more specific and so I treat it here as a containing inferable because the less specific *Dzigintsuan* is found within the larger noun phrase. The next example is similar, immediately following a reference to the *gamp* ‘village’:

- (53) B-40 *Da rīb gamp igi itsaŋa ba.*
 and 3pl village there R-see come
 ‘The people of the village saw (them) coming.’

Another example of a containing inferable is *Augustin bajin* ‘Augustin’s hand’, mentioned in section 4.4.1 above, since *Augustin* is contained within the whole noun phrase. Like brand-new anchored referents, the referring expressions of I_C entities are more complex, requiring full NPs, since the given link must be contained within them.

4.4.2.6 Examples of Evoked Referents

Evoked (discourse old) referents are the most numerous and generally the easiest to identify. Often they are pronouns, as in (51), above: *Da araja ba munti da ifinti siŋan* ‘He

⁹ The noun phrase *garam Dzigintsuan*, however, is part of the larger noun phrase *garam Dzigintsuan gampan*, and so I do not include the shorter term in the counts, as noted in the methodology section. It would not seem inappropriate to include it as long as we treat the data consistently. Here, however, we will follow Prince’s (1981) convention of treating these embedded referents as simply part of the descriptions of the referents they modify.

came and stood and put a ritual blessing on the spear'. Sometimes, however, they do not differ at all from their first use, as with proper nouns or the following example:

- (54) A-5,6 *Da wa' dzar' iwa? Ini ntaŋa dzar'...*
 and (um) pig R-come.out R-want spear-NMZ pig
 'And a pig came out. He wanted to spear the pig...'

The second use of *dzar'* 'pig' is evoked, while the first use is brand-new. Note how the choice of English articles in the translation reflects the changing status of the referent.

Evoked entities can be full NPs, proper nouns, or pronouns.

4.4.2.7 Examples of Situationally Evoked Referents

Situationally evoked entities are those that are evoked not from previous use in the text itself, but from the actual speech situation in which they occur. Again, these can be first and second person pronouns or demonstratives and demonstrative pronouns such as *this* and *that*, which point to a referent that is salient within the speech situation. I will consider subsequent occurrences of the same referent to be textually evoked rather than situationally evoked.¹⁰ Following are two examples of situationally evoked referents:

¹⁰ It is evident from Prince's examples that she considers only the first occurrence of a first or second person pronoun to be situationally evoked and subsequent occurrences to be textually evoked (1981:239). This is an issue with which I had to wrestle. On the one hand, any time a deictic pronoun like *you* or *I* is used, it takes its meaning from the real life context, whether or not it was previously used. In a conversation involving multiple parties, these pronouns change referents with every turn. Thus I can make a case for considering every instance of the Adzera pronoun *dzi* 'I' to be situationally evoked. On the other hand, after the speaker introduces himself as a participant in the discourse, subsequent mention of himself is both situationally and textually evoked. Though the speaker may be no more salient to his audience than he was before, he is now a given participant in the story by virtue of previous mention in the discourse. Since such referents are both situationally and textually evoked, it makes sense to use the default classification Evoked (E), so I follow Prince's convention of using E_S for the first mention only. However, when the first person plural pronoun *aga* 'we (exclusive)' is used, identification of the whole group—if not present—depends not only on the speech situation (who is talking), but also on the previous discourse in which the speaker specified the other members of the group. So even the first use of *aga* 'we (exclusive)' would be textually evoked in text A, where it refers to the speaker and his father.

- (55) A-3 *Itajin gum mision rai da iba ani...*
 R-leave work mission away and R-come here
 'He left mission work and came here...'
- (56) A-22 *Ramaj' iba ya iba yu dzi imin' gamp ani.*
 father-1pss R-come and R-come get 1sg R-from village here
 'My father came and took me from this village/here.'

In the first example *ani* 'here' takes its reference from the location in which the speech occurred, and it is also the first unambiguous reference to that place, thus it is situationally evoked. The second sentence does not contain such a clear example of a purely situationally evoked referent, but it is a good example to examine, nonetheless. In it there are two referents that are noteworthy. The first is *gamp ani* 'this village', which though situationally evoked, is also textually evoked (from the first example), so will be classified as such. The second is *dzi* 'me', which I am treating as situationally evoked, because it is the first "explicit" reference to the speaker himself and the first time the speaker enters the story as a character. However, since the beginning of the story, the speaker has used the inalienably possessed form *ramaj'* 'my father'—which has mandatory marking for possession. One can make a case that sentence A-22 above does not contain the first mention of the speaker, and thus the referent is also textually evoked. But I believe my classification of E_S is justified for the following reasons: (1) the speaker's options for referring to his father are limited and mandate a first person suffix; (2) the speaker does not use the first person pronoun *dzi* in front of *ramaj'* or anywhere in the preceding text; (3) the speaker does not use the longer possessive form *ramajganj'* 'my father' that would emphasize the possessor more; (4) as the preceding two observations indicate, the speaker wants to focus on his father, not on himself,

so the term *ramarj'* would be equivalent to the English terms 'dad' or 'daddy'; and (5) as mentioned, the speaker himself does not play a role in the story until this point. Ultimately, whether we classify the referent as E or E_s, he still has the same basic given status.

The examples above are the only two I found in these Adzera texts. In addition to personal and demonstrative pronouns, the E_s category can contain full NPs with proximal demonstratives, such as *gamp ani* 'this village' in (56), even though this particular NP was also textually evoked in this context.

4.4.3 Referents Within Quotations

Before we examine the distribution of these classifications, I want to discuss the classification of referents within quoted material. Embedded quotes within a narrative are like mini discourses of their own, and could be extracted and analyzed separately. If we did that, then the first *I*, *you*, or *we*, for example, of the quoted material would be situationally instead of textually evoked, and some of the other referents would be new rather than evoked. Consider (57), for example, with an embedded quotation:

(57) B-41 *Riara igi itsaja ba da iyai' gin,*
 Riara there R-see come and R-cry.out 3obj

«*Ayoe! Maramai' irut finin ifa ntar' fugai*
 EXCL Maramai R-with wife-3pss R-go spear crocodile

ya itarij' iba, ifits iba.>>
 and R-carry R-come R-carry R-come

'[The people of] Riara saw them coming and cried out about it,
 "Ayoe! Maramai went with his wife and speared a crocodile, and
 they carried it back.'"

In the text the referents in this quote are all evoked. However, in the original context in which this quotation was uttered, Maramai and his wife would have been unused referents, and the crocodile would have been a brand-new referent. This is somewhat reflected in the forms of reference used; none of them is pronominalized or ellipted. In another quote, the characters Kakarak and Ati refer to Maramai as *agi numuntuga' araga* 'our (INCL.) brother-in-law over there' the first time they see him. He is new (unused) to them, but old in the discourse. Other times quotes can introduce discourse-new referents, such as *Buarun'* (people/place name) in the first quotation of text A. These examples demonstrate that forms of reference within quotations can be somewhat skewed to reflect the given-new status in the minds of the original hearers of the quoted utterances. However, it is the status in the minds of the storyteller's audience that is important when analyzing texts with embedded speech. Therefore, all of these referents must be classified as evoked.

4.4.4 Quantitative Distribution of Referent Classes

In this section I will present the distribution of referent classifications in the Adzera texts in an attempt to uncover general patterns of the language. Prince 1981 makes a distinction between those referents that occur as subjects and those that occur as non-subjects. I divide the referents in the same way, but I also see the animate-inanimate distinction as an important one that also affects the distribution of referents in subject and non-subject positions. Therefore, I present separate breakdowns of the animate and inanimate referents for each text in addition to the total counts, which Prince did not.

Before we look at the distribution of classes, let us look at the overall patterns of the texts. In the two texts examined I classified 266 referents. Table 6, below, shows their general distribution in terms of animacy and status as grammatical subject/non-subject. The reader will note that text A contains many more overt referents than text B, but that the ratio of animate referents to inanimate referents in each text is roughly the same (about 2:1), and the distribution of subject versus non-subject referents for animate, inanimate, and total referents are remarkably similar for both texts. It is logical that the majority of animate

Table 6. General Distribution of Referents

	Text A	Text B
Total referents (266)	174	92
% of total referents that are Subjects	45%	40%
% of total referents that are Non-subject	55%	60%
Animate referents (173)	115 (66%)	58 (63%)
% of animate referents that are Subjects	63%	60%
% of animate referents that are Non-subjects	37%	40%
Inanimate referents (93)	59 (34%)	34 (37%)
% of inanimate referents that are Subjects	10%	6%
% of inanimate referents that are Non-subjects	90%	94%

referents and very few inanimate referents are subjects, since inanimate referents usually are not agents but subjects frequently are agents. Adzera's lack of a passive construction magnifies this tendency, since inanimate patients are not allowed as subjects unless the verb is intransitive and requires a patient or experiencer as the subject, rather than an agent. What may surprise some readers is that more of the referents are non-subjects than subjects. Most English speakers would expect all sentences to have subjects, but not all to have objects or

obliques. They would also expect most of the subjects to be referential, giving a higher subject count. As we saw in section 2.5.2, however, Adzera is a pro-drop language that allows subjects to be ellipled as long as they are recoverable from context. This is so frequent in Adzera, as we will see in chapter 5, that the distribution of overt subjects is drastically different than what we would expect in English.

Beginning on page 87, I present distribution tables for total referents, animate referents, and inanimate referents. Tables 7 and 8 present the total referent counts for texts A and B, respectively. There is more variation between texts in these two tables than in table 6, above, and yet the same general patterns hold even though the percentages are different. Note that the majority of both subjects and non-subjects in each table are evoked, and that there are many more new non-subjects in each table than new subjects. These results fit the expectation that more referents will be introduced to a discourse in positions or roles other than the subject, since given entities frequently fill the subject slot in SVO languages.

Looking at animate referents in tables 9 and 10 (page 88), one finds the distribution to be similar to that of all referents, but with a smaller percentage of inferred referents and a higher percentage of evoked entities, especially in the non-subject columns. Notice also that subjects outnumber non-subjects among animate entities.

The counts for inanimate referents given in tables 11 and 12 (page 89) offer even more insight. In particular, no new entities are introduced as subjects, while evoked and inferable entities do occur as subjects, suggesting that the latter two have something in common. As mentioned in chapter 3, Birner (1994, 1997) provides evidence that inferable and evoked referents have a similar distribution and are treated alike under certain

circumstances. We also see in English the definite article used with inferable entities and evoked entities alike. These last two tables support the claim that inferable entities pattern like evoked entities by showing the same kind of distribution, in stark contrast to the new entities which never occur as subjects. The distribution of only eight inanimate subjects is hardly conclusive evidence, but it does offer support for other similar findings.

It is noteworthy that the inanimate referent tables for the two texts are nearly identical when one compares the percentage totals for each of the three broad categories—evoked, inferable, and new. The totals in the subject columns are, in fact, identical, and the non-subject columns are very similar. In the animate tables, however, there is much more variation. Text A has significantly more new referents than text B, and, therefore, a smaller percentage of evoked referents. Some of these extra new referents in text A come from the list of groups who were baptized, but there are also more simply because there are more characters throughout the text. Inanimate entities, on the other hand, serve as props and tend to be more evenly distributed.

Table 7. Total Referents in Text A
(by number and percentage)

	Subjects (79)		Non-subjects (95)	
Evoked				
E	64	81.0%	59	62.1%
E _s	0	0.0%	2	2.1%
Total	64	81.0%	61	64.2%
Inferable				
I	6	7.6%	8	8.4%
I _c	4	5.1%	3	3.2%
Total	10	12.7%	11	11.6%
New				
U	2	2.5%	19	20.0%
BN _A	0	0.0%	1	1.1%
BN	3	3.8%	3	3.2%
Total	5	6.3%	23	24.2%

Table 8. Total Referents in Text B
(by number and percentage)

	Subjects (37)		Non-subjects (55)	
Evoked				
E	34	91.9%	39	70.9%
E _s	0	0.0%	0	0.0%
Total	34	91.9%	39	70.9%
Inferable				
I	1	2.7%	8	14.5%
I _c	1	2.7%	0	0.0%
Total	2	5.4%	8	14.5%
New				
U	0	0.0%	1	1.8%
BN _A	1	2.7%	3	5.5%
BN	0	0.0%	4	7.3%
Total	1	2.7%	8	14.5%

Table 9. Animate Referents in Text A
(by number and percentage)

	Subjects (73)		Non-subjects (42)	
Evoked				
E	61	83.6%	30	71.4%
E _s	0	0.0%	1	2.4%
Total	61	83.6%	31	73.8%
Inferable				
I	5	6.8%	1	2.4%
I _c	2	2.7%	0	0.0%
Total	7	9.6%	1	2.4%
New				
U	2	2.7%	9	21.4%
BN _A	0	0.0%	1	2.4%
BN	3	4.1%	0	0.0%
Total	5	6.8%	10	23.8%

Table 10. Animate Referents in Text B
(by number and percentage)

	Subjects (35)		Non-subjects (23)	
Evoked				
E	33	94.3%	19	82.6%
E _s	0	0.0%	0	0.0%
Total	33	94.3%	19	82.6%
Inferable				
I	0	0.0%	1	4.3%
I _c	1	2.9%	0	0.0%
Total	1	2.9%	1	4.3%
New				
U	0	0.0%	0	0.0%
BN _A	1	2.9%	3	13.0%
BN	0	0.0%	0	0.0%
Total	1	2.9%	3	13.0%

Table 11. Inanimate Referents in Text A
(by number and percentage)

	Subjects (6)		Non-subjects (53)	
Evoked				
E	3	50.0%	29	54.7%
E _s	0	0.0%	1	1.9%
Total	3	50.0%	30	56.6%
Inferable				
I	1	16.7%	7	13.2%
I _c	2	33.3%	3	5.7%
Total	3	50.0%	10	18.9%
New				
U	0	0.0%	10	18.9%
BN _A	0	0.0%	0	0.0%
BN	0	0.0%	3	5.7%
Total	0	0.0%	13	24.5%

Table 12. Inanimate Referents in Text B
(by number and percentage)

	Subjects (2)		Non-subjects (32)	
Evoked				
E	1	50.0%	20	62.5%
E _s	0	0.0%	0	0.0%
Total	1	50.0%	20	62.5%
Inferable				
I	0	0.0%	7	21.9%
I _c	1	50.0%	0	0.0%
Total	1	50.0%	7	21.9%
New				
U	0	0.0%	1	3.1%
BN _A	0	0.0%	0	0.0%
BN	0	0.0%	4	12.5%
Total	0	0.0%	5	15.6%

4.5 Summary

Prince's taxonomy of assumed familiarity, as the name suggests, is based not on the hearer's actual familiarity with a referent, but on the speaker's assumptions about the hearer's familiarity. Though the hearer's knowledge would be impossible to determine, the forms of reference that the speaker uses can tell us a great deal about the assumptions he is making. These assumptions can sometimes lead to miscommunication or confusion if the speaker assumes too much knowledge or inadvertently leaves out certain details that the hearer requires. There were examples that seemed to demonstrate this in one of the Adzera texts, but a more likely interpretation is that the references represent irrelevant information as part of a strategy to compensate for the lack of passives in the language. Whether or not to include the vague third person pronouns in question is open for debate, but I have included them in this analysis. Because the forms of reference do not always line up with what one would expect and because every analyst has different assumptions, the method is not totally objective and requires some intuitive thought. Nevertheless it can offer great insights when applied consistently to enough data.

Classifying all the referents in the Adzera texts and counting their distribution yielded interesting results. Based on the analysis of these texts I make the following generalizations about Adzera oral narratives. Animate entities tend to outnumber inanimate entities (roughly 2:1). Overall there are fewer overt subjects than overt non-subjects, primarily due to subject ellipsis. Among animate referents overt subjects do outnumber non-subjects (about 3:2). Inanimate entities rarely occur as subjects, and when they do they are never new. The vast majority of subjects are evoked entities. Speakers tend to introduce new referents in non-

subject positions, but they do sometimes introduce new animate referents (1 out of 3 or 4) in subject position. This is likely due at least in part to Adzera's lack of existential or presentational "there" devices (cf. Ward 1999) or dummy subjects. At least where inanimate entities are concerned, inferable referents have a similar distribution to evoked referents, offering support to claims that they pattern the same. Texts may differ considerably in the number of characters and thus the number of new animate referents, but may have remarkably similar inanimate counts, suggesting that props tend to be more evenly distributed than participants. Finally, the reader can see that the animate-inanimate distinction is an important one to make in this kind of analysis.

CHAPTER 5

A QUANTITATIVE ANALYSIS OF TOPIC/PARTICIPANT CONTINUITY

5.1 Introduction

Chapter 3 included a brief overview of Givón's (1983, 1995) quantitative approach to ¹ topic/participant continuity. This chapter will apply his approach to the Adzera texts, as a way to track "topics" or referents, especially animate participants. In the last chapter we examined overt referring expressions and, based on their forms, classified their respective referents according to the amount of familiarity the speaker was assuming or appeared to be assuming the hearer to have. While the previous chapter was limited to overt forms of reference, this present chapter accounts for the ellipted referents so common in Adzera. It was helpful in the previous chapter to look at the less flexible distribution of inanimate referents, which allowed us to see a similar pattern between evoked and inferable entities. In this chapter we will look at both animate and inanimate referents, concentrating mostly on the former. The analysis will seek to quantitatively measure topic/participant importance, reflected by persistence, and topic accessibility, which is partially determined by referential distance and potential inference. The latter of the two, accessibility, is of most interest to us in this paper. This approach will reveal patterns about the tracking of referents in the Adzera language, which should help to explain the forms of reference that were used in the previous chapter and their distribution. First I will cover the methodology I used, and then the results of the analysis.

5.2 Methodology

5.2.1 Measurements

Chapter 3 introduced three measurements of topic continuity proposed by Givón (1983, 1995): referential distance, potential interference, and topic persistence. The first two measure topic accessibility or availability and the last one measures longevity. This section briefly reviews these measurements and defines more specifically how they are measured in this paper.

Referential distance (RD) measures the absence, if any, of a referent from the discourse by counting the number of clauses from the occurrence in question to the most recent previous occurrence of the referent (explicit or implicit). The value is simply the number of clauses to the left (up to a set maximum). Low values (little or no gap) indicate accessible, or easily identifiable, topics, while high values (large gap or no previous mention) indicate inaccessible topics. The minimum value is 1 (most accessible/continuous), and I follow Givón 1983 in setting the maximum value to the somewhat arbitrary figure of 20 (least accessible/continuous). The first mention of a topic is obviously the most difficult to identify, and therefore we must assign the maximum value of 20; we can think of this as an infinite gap or absence in the previous discourse.

Potential interference (PI) measures the potential for ambiguity caused by interference from other referents in the immediate context of the referent in question. Other referents in the immediate discourse can make it difficult for the hearer to positively identify which referent the speaker has in mind if the speaker is not careful to specify. Again, the relevant question we asked in chapter 3 was: "If the minimum possible markings were used,

could the referent be confused with another referent with similar semantic attributes in the immediate context?" If the answer is yes, then I assign the value 2; otherwise, I assign the value 1. Different researchers have used different numbering systems to measure PI, but this is the most common one used by contributors to Givón 1983. How one defines the "immediate context" is also up to the researcher, but most have considered it to be three clauses to the left, so I also set the context at three clauses to the left. I also assign the value 2 to all first mentions of an entity in a discourse following the convention in Givón 1983.

Topic persistence (TP) measures the longevity of a topic, which reflects its importance in the discourse. Instead of looking back at previous clauses as the other two measurements do, TP looks ahead to the subsequent clauses of the discourse to see how long a topic remains salient. Givón 1983 defined persistence as the number of uninterrupted clauses to the right that the topic remains "a semantic argument of the clause" in any role (1983:15). The main problem inherent with this method of measuring persistence is that counting uninterrupted clauses does not allow for even a small gap in a string of references to an otherwise highly persistent topic. Parenthetical explanations or other embedded material can too easily wreak havoc with persistence measurements. Therefore, in later work Givón eliminates this problem by modifying his counting methods for topic persistence. Instead of the number of clauses immediately to the right, he counts the number of references within the next ten clauses (1995:79). This study employs this revised counting method. The minimum value is 0 for a topic that is not mentioned again in the next ten clauses, and there is no

absolute maximum, though there are natural limitations on how many times a referent can be mentioned in ten clauses.¹

Table 13 below summarizes the three measurements used in this chapter, giving the range and significance of their values.

Table 13. Interpreting the Measurements

	SCALE	LOW VALUE	HIGH VALUE
RD	1-20	continuous/accessible topic	discontinuous/inaccessible topic
PI	1-2	sometimes can be left implicit	may need more explicit reference
TP	0+	discontinuous topic	continuous topic

The reader should note that while low values for the first two measurements indicate more easily identifiable (and, therefore, more continuous) topics, a low value for the third measurement, TP, indicates a discontinuous topic, just the opposite. Also note that when both RD and PI are low, a speaker will often use the minimum form to designate a referent, but if either measurement is high, a speaker will often use a more explicit reference than the other measurement might predict. So, for example, if a highly salient participant has been mentioned in the previous clause (RD = 1), one would expect the minimum of marking for that referent, such as a pronoun in English or ellipsis in Adzera, but if there is another character mentioned in the immediate three clauses to the left (PI = 2), a more explicit reference may be required instead. It is in this way that these two measurements work together to determine the identifiability of a topic and the form it will take.

¹ If, on the other hand, we count the number of those ten clauses in which the referent occurs, the maximum would be ten. This would make the counting task easier, but it may not be a better measurement of persistence. I do not believe that Givón sets a maximum of ten. But even if he does, I do not, allowing instead for more than one reference within a clause or what is counted as a clause (to be discussed below).

5.2.2 Counting Clauses

I divided the texts into clauses, but the division was not as straightforward as one might think. The non-linear nature of natural discourse poses some problems. Natural speech has different levels with some elements embedded within others.

Direct quotations are an example. They break up the chronological flow of a narrative by inserting what a character said at a given point in time. Quotations often do not move the event line of a narrative forward, yet they can provide valuable information that helps the hearer to understand subsequent events. Sometimes they contain a retelling of past events in the discourse; other times they are a statement of a character's intentions rather than what actually happens. They may contain commands, warnings, threats, predictions, desires, and more. A direct quote may or may not contain a reference to the participants that are found before or after it on the event line. Should they be included or excluded? The topics within quotations often may not continue to be topics outside the mini discourses of the quotations themselves. Therefore, in contrast to the approach in chapter 4, I do not measure the continuity of occurrences inside quotations. Furthermore, following the convention Givón uses in his analysis of Ute (1983:157), non-appearance inside a direct quotation does not count as a gap, but appearance within the quotation does count. To achieve this without further complicating the counting methods, I include a direct quote as part of the clause in which it is embedded. So if a direct quote happens to be two clauses long, those two clauses plus the one in which they are embedded count as only one clause.

Like Givón (1983:157) I treat relative clauses the same way as quotations, since they are embedded in noun phrases, but I extend the same treatment to complement clauses as

well (which he does not), since they are embedded in verb phrases and seem to function much in the same way that direct quotations do. Therefore, a main clause with a relative clause or a complement clause (or both) counts as only one clause for counting purposes. All other clause types are considered separate clauses and are treated the same as main clauses.²

A clause can be as little as a single verb separated from other clauses by conjunctions or punctuation; however, a string of serial verbs (covered in section 2.5.3) counts as only one clause, as it effectively communicates only one action or event. After deciding, according to the above criteria, what constitutes a “clause” for our purposes, I numbered each clause and listed it on a separate line with embedded material tabbed in.

5.2.3 Defining Topics/Participants and Counting Their Occurrences

A topic can be a noun phrase or its substitute occurring anywhere in a sentence. As in chapter 4, only those entities that are referential count as topics. Thus a reference to pigs in general or a hypothetical pig does not constitute a topic/participant, because no specific referent or example of the type “pig” is in mind. However, if a particular pig or group of pigs is in mind, I count it as a referent. I also exclude propositional information as much as possible. However, the line between a referent and a proposition becomes fuzzy at times. Consider, for example, the following sentence that appeared as example (57) in chapter 4, repeated here as example (58):

² There may be many other valid (perhaps better) ways of dividing the text and counting clauses; however, none of them will be without its problems when it comes to trying to quantify continuity and also take into account the syntactic and discourse-level hierarchies of speech. I attempt to apply the method I have chosen consistently, which, I believe, is the key.

(58) B-41 *Riara igi itsarja ba da iyai' gin,*
 Riara there R-see come and R-cry.out 3obj

<<Ayoee! *Maramai' irut finin ifa ntay' fugai*
 EXCL Maramai R-with wife-3pss R-go spear crocodile

ya itariy' iba, ifits iba.>>
 and R-carry R-come R-carry R-come

'[The people of] Riara saw them coming and cried out about it,
 "Ayoee! Maramai went with his wife and speared a crocodile, and
 they carried it back.'"

The Adzera pronoun *gin* '(about) it' is always used to refer in the third person to people, things, or ideas that have just been mentioned. It always has an antecedent in the immediately preceding context and commonly occurs as a resumptive pronoun in relative clauses. In this case the antecedent is taken to be what they saw, namely that Maramai and his wife were coming along carrying a dead crocodile. It is not just Maramai, his wife, or the crocodile that they are crying out about, but all of it together and the inferred proposition that Maramai killed a crocodile and they carried it back, which is reflected by what they say. Therefore, I do not count *gin* in this instance as a topic, since it refers to something that is propositional in nature.

Following Givón 1983 I count both explicit and implicit occurrences of a referent. This includes any argument that is semantically implied by the verb, whether or not there is any overt marking for it anywhere; for example, in the first line of (58), above, there is an implied argument between 'see' and 'come'. This must be counted. However, I count the subject of a serial verb sequence only once rather than once for each individual verb. This is illustrated in example (59):

- (59) B-9 (Clause 23) *ya* \emptyset *irua'* *iru'* *ifan*
 and R-fall R-go.down R-go
- (Clause 24) *ya* \emptyset *itip* *ifa* *yu* *a'* *puatsi* *iba*
 and R-repeat R-go get (uh) grass R-come
- (Clause 25) *ya* \emptyset *iba* *iyab*
 and R-come R-go.up
 '...and she went down and brought back some more grass and came back up...'

This comes from the seven-clause crocodile-nest-building sequence that we saw in section 2.5.3. The conjunctions divide the action sequence into clauses, but, as shown earlier, within each clause is a set of serial verbs that function as one. There are nine verbs in these three clauses, but because I consider each verb series as a single verb, I count only three references to the crocodile as subject, instead of nine. Similarly, I count each instance of possession as only one reference, even if there is more than one marking for it as in example (60):

- (60) A-49 *Gurump narung-gan...*
Gurump child-3pss
 'Gurump's son'

Both the name *Gurump* and the possessed noun *narung-gan* 'his son/child' might be said to contain a reference to the same participant, and the latter could be used without the proper noun preceding it when the possessor is clear from context. The second marker is simply third person agreement on the inalienably possessed noun (cf. section 2.4.3), which is required even when a more explicit reference to the possessor occurs, as it does in this case. This is really only one reference to *Gurump*, and is counted as such. Note, however, in the methodology of Prince in the previous chapter, I did not count the reference to *Gurump* in this instance since it is embedded inside the referring expression of another referent.

Following Givón's methodology, however, I do count these types of occurrences in this chapter.

The reader can see that there is not a one-to-one correspondence between referents counted in this method and what was counted in Prince's method. So far we have seen three differences in this method, namely the inclusion of zero anaphora, the exclusion of referents within quotations, and the inclusion of referents embedded within the referring expressions of others. There is one other type of referent that I have excluded, that of the introduction of a person's name. While such examples are evidently important to Prince's method, which looks at new referent introduction and the given-new status of referents in general, Givón's method tracks continuity and is applied best to sequentially-ordered narrative. That is not to say that there are no other breaks in the action in which the narrator gives further explanation, and within which I count referents. But the type of formulaic equation used for introducing names presents a challenge for the analyst employing this method. Like English, the Adzera formula for name introduction is as follows: 'My name (=) Gadan' or 'Gurump's son's name (=) Mugusa'. First of all, 'my name' represents an inanimate characteristic of an animate referent. Even though the character can be very salient, this inanimate characteristic of the participant appears for the first time in the text and often the last. Do we assign the reference the RD value of 20 because the name has not been mentioned? It is not a part of the animate referent, such as an arm, that we can easily separate from the character in our minds and that might receive special attention in a text. Rather as a representation of the identity of the character, it is a further description of the character. Names seem to fall into a class of their own, and blur the distinction between animate and inanimate referents. What

do we do with the name itself, on the “other side” of the zero copula? Do we count it as a separate reference? I would argue that it is part of the same reference by virtue of the equative nature of the sentence. Does it count as a reference to the animate participant? Consider the sentence sequence from text A (sentences 48-49), glossed here in English: ‘Then Mugusa came running. Gurump’s son’s name [was] Mugusa.’. Is the second *Mugusa* a reference to the first? It certainly is, and yet it is equated with his name. Is ‘Gurump’s son’s name’ a separate entity from Mugusa himself? The reader can easily see the problems with categorizing these names as animate or inanimate referents and assigning RD, PI, and TP values to them. Therefore, I do not treat a phrase like ‘Gurump’s son’s name’, as a separate referent to be counted and scored, but I do count the references to the animate participants found within such phrases (i.e. ‘Gurump’ or ‘Gurump’s son’). Because the methods for counting occurrences of discourse referents vary in several ways between the two approaches, the referent counts in this chapter do not match those of the previous chapter. Table 14 summarizes the differences between the two approaches employed in this study, in terms of what is counted:

Table 14. Contrasting the Two Approaches

Assumed Familiarity (Prince) – Chapter 4	Topic Continuity (Givón) – Chapter 5
analyzes overt references only	analyzes zero anaphora also
analyzes references within quotations	does <u>not</u> analyze references within quotations but counts them when measuring the continuity of other references
does <u>not</u> analyze references embedded within other referring expressions	analyzes references embedded within other referring expressions
analyzes name formulas	does <u>not</u> analyze name formulas themselves, but does analyze references within them

5.2.4 Counting and Computing Methods

I entered each occurrence of a referent into a spreadsheet program along with its clause number and grammatical position (i.e. subject, direct object, etc.). I made separate lists for animate and inanimate topics, and within each of these I divided them according to their morphological manifestation as full noun phrases, pronouns, zero anaphora (ellipsis), or possessor agreement only (as discussed above). Within each list I made a column for each of the three measurements, RD, PI, and TP. Then I pored through the texts one at a time, manually performing the required counts for each measurement and assigning the appropriate values to each occurrence. The spreadsheet calculated the average values for each measurement for full NPs, pronouns, ellipsis, and possessor agreement. These average values are what Givón 1983 compared, and they make quick comparisons easy. But Givón 1995 uses a percentage method for representing the results of the counts, by designating a high-low continuity cut-off for each measurement and giving the numbers and percentages of referents that fell above and below that cut-off. While I use both methods in this study, I use the latter method more to present the data. While the second method requires more manual calculation and the results are not always immediately obvious at first glance, it helps to illuminate the data and reveal patterns not previously seen. Furthermore, it helps to minimize the effect of the arbitrary maximum for RD (20), which can easily skew the averages, especially in small samples.

I divided the data in different ways to illuminate patterns related to different factors. In addition to animacy and forms of reference, I examined subjects vs. non-subjects and main characters vs. supporting characters. As we will see below, the results show that there is a

relationship not only between continuity and the form of reference employed, but also between continuity and grammatical position, continuity and animacy, and grammatical position and animacy. So there are several factors to be considered in relation to how topics/participants are expressed.

5.2.5 Illustration of the Methodology on a Portion of Text

In chapter 4 I illustrated the approach using a portion of text. Here I will illustrate the topic continuity approach by applying it to the same portion of text. This will enable the reader to see what each method includes in the same small portion of text. Since zero anaphora are counted in this method, I represent them in the text with the null symbol (\emptyset), and I number each "counted" or measured reference to a discourse entity with a numerical subscript, as I did in chapter 4. The text is presented in example (61):

- (61) A-48 *Da Mugusa₁ irunt iba.*
 and Mugusa R-run R-come
 'Then Mugusa came running.'
- A-49 [*Gurump₂ narujgan*]₃ *bijangan Mugusa₃.*
 Gurump child-3pss name-3pss Mugusa
 'Gurump's son's name was Mugusa.'
- A-50 *Da \emptyset_4 isalim arigi₅, \emptyset_6 ini bina₇.*
 and R-send that.one R-say like.this
 '(He) sent him, saying,'
- A-51 <<*Warunt da wani da mision*
 IMP-run and IMP-say to missionary
i ya' Sinkom buy' a'... nabia₇
 Comp (uh) Sinkom (FUT) (IR-) FUT-come

nais a', mision da naga Ø
 FUT-kill (uh) missionary and FUT-eat

da gubu' arani.
 TIME day this.one

'<<Run and tell the missionary that Sinkom will come and kill him and eat (him) today.'

A-52 <<*Da* Ø *natip rungan*
 and.so HORT-prepare self-3pss
ma Ø *nafan ya* Ø *nagup.>>*
 or HORT-go and HORT-hide

'<<So (he) should get himself ready or go and hide.>>'

A-53 *Da ramaj'7 ini da dzi8,*
 and father-1poss R-say to 1sg

'And my father said to me, ...'

I consider the rather lengthy direct quotation in A-51 and A-52 to be embedded within the clause that introduces it, namely the second clause of A-50, and I have attempted to show this relationship by indenting the quotation. The disadvantage of this treatment is that the analyst must consider a lengthy section of text (8 clauses) as one clause, which can cause higher persistence scores, since what is counted as ten clauses may actually be seventeen. The advantage is that direct speech, which may not contain reference to the salient topics outside of the quotations, is not allowed to interrupt the continuity of those topics. The reader will notice that none of the references within direct speech is marked for inclusion in the analysis. These references do, however, count when measuring the persistence and referential distance of topics outside of the quotation. For example, there are seven references to the main character in A-51 and A-52 that figure in the topic persistence counts of the last measured reference to the main character. They also count in the RD

measurement for *ramanj* 'my father' in A-53, giving a value of 1. Even without analyzing the references within the quotations, this portion of text includes eight references to be measured, including zero anaphora. The reader will recall that the approach in chapter 4 yielded ten references, including those inside the quotation but not including zero anaphora (as in A-50) or references embedded within other referring expressions (as in A-49).

Now I will take a closer look at each reference. The first is *Mugusa*. Since this is the first mention of the referent, it receives an RD value of 20 and a PI value of 2. Since I considered the copula construction in A-49 to contain just one reference to *Mugusa*, the reference in A-48 has a TP value of 4 (the copula construction in A-49, *arigi* in A-50, and two implied references in the imperative clauses in A-51).

The second reference *Gurump*, has an RD value of 4 (last mention in A-41 is only four "clauses" away when quotations are considered to be embedded), a PI value of 2 because other human referents are mentioned within the previous three clauses (including the main character and the narrator), and a TP value of 2 as the implied subject of both clauses in A-50.

The third reference, 'Gurump's son/*Mugusa*', has an RD value of 1 (being introduced in the previous clause), a PI value of 2 (due to the presence of other characters in the preceding clauses), and a TP value of 3 (one less than the previous occurrence of the referent).

I will skip the fourth reference for a moment and look at the fifth first. The demonstrative pronoun *arigi* 'that one/him' is also a reference to *Mugusa*. It has RD and PI values of 1, because *Mugusa* is mentioned in the previous clause and there are no other

characters besides Gurump and the narrator himself in the preceding three clauses. Gurump does not present potential interference because it is clear from context who came running and who must have sent whom, even if this reference were a zero anaphor. The TP value of this reference is 2, because the only mentions in the next ten clauses are the two implied references from the imperative verbs in the quotation.

In numbers 4 and 6 (in A-50), Gurump is the implied subject of both verbs. RD and PI are 1 in both cases, the same reason as above applying regarding the lack of PI. Number 4 has a TP value of 1 (i.e. counting #6 as the last reference), and number 6 has a TP value of 0 (no subsequent mention in the ten clauses, or even the rest of the story).

References 7 and 8 are found in A-53. Both have PI values of 2, because of potential interference from Gurump, Mugusa, and Sinkom. Number 7, *ramay'* 'my father', has an RD value of 1 because of the seven references to him within the quote, as noted above. The last prior reference to the narrator (not counting the possessor marking of #7) comes five "clauses" before (not shown in sample), so the RD is 5 for reference #8. Though it is not shown in the sample, I calculated TP values for #7 and #8 to be 8 and 10, respectively.

It should be noted that the TP values for the last two references in this small sample of text cannot be determined from only the portion of text presented in (61), above. This is technically true of the other six references as well, but because the others decayed within the text portion shown, the TP values correspond exactly with the subsequent appearances shown in the sample. Table 15 summarizes the referents analyzed in this portion of text, and their RD, PI, and TP values:

Table 15. Summary of Sample References and Their Continuity Scores

#	NP	RD	PI	TP
1	<i>Mugusa</i>	20	2	4
2	<i>Gurump</i>	4	2	2
3	<i>Gurump narungan / Mugusa</i>	1	2	3
4	∅ (Gurump)	1	1	1
5	<i>arigi</i> (Mugusa)	1	1	2
6	∅ (Gurump)	1	1	0
7	<i>ramanj'</i>	1	2	8
8	<i>dzi</i> (narrator)	5	2	10

In my actual calculations I listed these references in different columns for each form of reference: full NPs, pronouns, etc. With the values listed in that way, it is easy enough to average each column or to take counts of RD values of 1 to 3 and those greater than 3, for example. Finally, the reader should note that the above passage was chosen for purposes of comparison with the approach in chapter 4, because it contained all of the features that distinguish the two methods. Because of this, it is not the easiest passage to analyze with this method. Givón's approach works better with straight sequential action narrative portions.

5.3 Predictions

Givón makes predictions based on his "scale of phonological size" illustrated in figure 4:

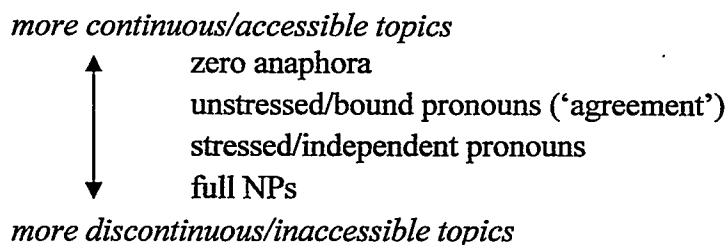


Figure 4. Givón's scale of phonological size (taken from Givón 1983:18).

The scale predicts that we should see, in any language that employs these grammatical forms, the same general order of forms corresponding to the topics they express from most continuous to least continuous. Givón expresses the principle behind this phenomenon in this way: “The more disruptive, surprising, discontinuous or hard to process a topic is, the more *coding material* must be assigned to it” (1983:18). This is thoroughly consistent with Grice’s maxims of quantity. This, of course, is not an absolute prediction; referents can take different forms than their accessibility scores predict, but statistically the smaller the phonological size, the greater the accessibility according to the measurements. Based on this we can predict that all of the above forms of reference that Adzera speakers have at their disposal will fall into that same general order with respect to topic accessibility. Clearly Adzera speakers use zero anaphora, unbound pronouns, and full noun phrases. Therefore, I can predict with a fair degree of confidence that topics marked in Adzera by zero anaphora will be more accessible than topics marked by pronouns, which will be more accessible than those marked by full noun phrases (including unmodified nouns and proper nouns in this analysis).

There is a fourth category that I measured in the Adzera texts that would seem to fit most naturally into Givón’s unstressed/bound pronoun (agreement) slot, namely possessor agreement. Even though Adzera verbs are not marked for subject (or object) agreement, inalienably possessed nouns must always be marked for possessor agreement and can occur with or without an explicit possessor (as discussed in section 2.4.3). In fact, this possessor agreement is mandatory on all nouns whenever a possessor is indicated, whether or not the possessor is made more explicit. If they occur without an explicit possessor (i.e. no pronoun

or noun phrase preceding the possessed noun), then the agreement is the only reference to the possessor, whose exact identity must be recoverable from context. In one way these are very much like instances of zero anaphora, yet there is an overt marking in the agreement.

Consider the noun phrases in table 16:

Table 16. Possessor Marking

Possessor Type	Possessor	Possessed Noun	Agreement
Full noun phrase	<i>Gurump</i> Gurump	<i>naruN</i> ³ child	- <i>gan</i> -3pss
Pronoun	<i>dzi</i> 1sg	<i>buN</i> in-law	- <i>ga'</i> -1pss
Zero/Agreement only	∅	<i>raiN</i> brother	- <i>gan</i> -3pss

All the examples in table 16 come from the texts. I included full noun phrase possessors in the full NP list and pronoun possessors in the pronoun list, but I created a fourth list for the third kind of possessor marked only by agreement. I would expect the references in this fourth category to have accessibility values close to those of the zero anaphora (ellipsis) list, since the ellipsis of the possessor before the possessed noun indicates a very salient referent. Since the agreement marking is an overt reference, it is reasonable to expect the scores to fall between those of ellipsed and pronominal forms, where Givón places unstressed/bound pronouns (agreement). On the other hand, since this is a different kind of agreement with a much more limited distribution than subject-verb agreement, it may not pattern exactly as one would expect of subject-verb agreement.

³ For simplicity here, N represents an unspecified nasal, whose manifestation depends on whether its possessor is first, second, or third person. In cases like these in which the longer forms of possession are used (-*gan'* or -*ga'*, -*gam*, -*gan*), the nasal assimilates to the [g], and the possessor agreement is expressed solely by the person morpheme joined to the end of the possessive morpheme -*ga*. See section 2.4.3 for more explanation.

Givón (1983) also puts forth other scales and predictions regarding continuity. He predicts that animate participants will be more continuous than inanimate topics, and that the topicality or continuity of semantic case roles can be expressed as a hierarchy: agents > dative/benefactive objects (recipients/beneficiaries) > accusative objects (patients) > others, where the symbol > means “is more continuous/accessible than” (1983:22). Similarly he presents “grammatical cases” or roles in the following hierarchy: subject > direct object > others (1983:22).⁴ The overlap of these scales and the reasons for it seems fairly intuitive. Animate beings tend to be the main participants in a narrative, while inanimate objects tend to be relatively unimportant and briefly mentioned props. Furthermore, subjects tend to be volitional agents, which are almost always animate, while inanimate objects are more often patients occurring as direct objects or fill some less important semantic role expressed as an oblique. The fact that we can talk about the relative importance of sentence constituents suggests that there is a logical hierarchy of importance in our minds that would confirm the subject of a sentence as typically the most important and topical entity. Of course none of these observations is a hard and fast rule; all of them are general tendencies that we can represent statistically. In the next section I will offer support for Givón’s predictions concerning animate vs. inanimate topics and the grammatical role hierarchy (i.e. subjects vs. objects, etc.).

⁴ Other researchers, however, have shown evidence that given information tends to appear in sentence-initial position and new information tends to appear in non-initial positions, regardless of the grammatical subject or agent of the sentence. This is particularly evident in instances of subject-predicate inversion, topicalization of a non-subject entity, passive constructions with *by*, and existential and presentational *there* constructions (Birner 1994, 1997; Birner and Ward 1998; Ward 1990, 1999). Therefore, the high topicality of subjects in SVO and SOV languages is, in part, a function of sentence position.

5.4 Results

5.4.1 Overview

In this section we will view the results for topic continuity measurements of both texts. The total number of references measured is 384—294 refer to animate referents and 90 to inanimate. Of these, 209 are subjects and 175 are non-subjects. Since there are many factors that we need to consider, I will attempt to present them in a logical and orderly manner. We will begin by examining how topic continuity values relate to phonological form, first in terms of accessibility, measured by referential distance (RD) and potential interference (PI), and then in terms of topic persistence (TP). The accessibility of referents will be our main focus, because accessibility ultimately determines the forms of reference that entities take. Then we will examine how subjects and non-subjects vary with respect to topic continuity and animacy.

5.4.2 Accessibility and Form

5.4.2.1 Accessibility of Animate Participants

I present here separate tables for text A and text B of accessibility of animate referents in relation to the phonological form of the references, from zero anaphora to full noun phrases. The ordering of forms in the tables reflects our expectations based on Givón's scale of phonological size.

Table 17, below, presents the average RD and PI values for text A. Remembering that low RD and PI values represent high accessibility, the reader should note that the average RD value for possessor agreement (3.06) seems out of place with respect to the average pronoun value (2.85). This figure is skewed by the fact that 23 of the 32 occurrences

Table 17. Average RD and PI values for Text A
(Animate Referents)

	N	RD	PI
Ellipsis	35	1.03	1.20
Poss. Agr.	32	3.06	1.06
Pronouns	27	2.85	1.59
Full NPs	57	7.12	1.89

have first person agreement, referring to the speaker himself who is always salient regardless of how recently the last mention occurred. We discussed this in the section of chapter 4 about situationally evoked entities. I argued there that *ramanʔ* ‘my father’ (which makes up 22 of the 23 first person occurrences) does not contain a proper reference to the narrator, but rather includes the agreement that is mandatory for inalienably possessed nouns, and that this is the minimum possible reference to his father that the speaker has at his disposal. To be certain, the narrator does not intend to bring himself to mind every time he mentions his father in the story, and these instances could be translated quite naturally as ‘Dad’. Similarly there is one occurrence of *rinanʔ* ‘my mother/Mom’. Therefore, from now on I will exclude these occurrences of first person agreement. Note how excluding them causes the RD values to line up nicely in table 18. With respect to RD, possessor agreement falls between ellipsis and pronouns, but in terms of PI, it appears to be more accessible than ellipsis.

Table 18. Adjusted Average RD and PI values for Text A
(Animate Referents)

	N	RD	PI
Ellipsis	35	1.03	1.20
Poss. Agr.	9	1.11	1.06
Pronouns	27	2.85	1.59
Full NPs	57	7.12	1.89

By setting numerical boundaries for high-low accessibility, one can represent these same results in terms of percentages of referents above and below the cut-offs. Following Givón 1995 I set the RD threshold at three clauses, so an RD value of 3 or less indicates high accessibility, while an RD value of greater than 3 indicates low accessibility. PI values are already binary; 1 indicates no interference (and thus more easily identified), and 2 indicates potential interference. Representing the results in this way gives a slightly different picture as shown in table 19:

Table 19. Percentages of High and Low RD and PI values for Text A
(Animate Referents)

	N	RD 1-3	RD > 3	PI = 1	PI = 2
Ellipsis	35	35 100%	0 0%	28 80%	7 20%
Poss. Agr.	9	9 100%	0 0%	8 89%	1 11%
Pronouns	27	23 85%	4 15%	11 41%	16 59%
Full NPs	58	36 62%	22 38%	6 10%	52 90%

Table 19 indicates that there is no difference in the RD values of referents expressed by ellipsis/zero anaphora and possessor agreement, though there is less potential interference with possessor agreement. This suggests that possessor agreement might encode more accessible topics than zero anaphora, but two things must be remembered. First, the occurrences of possessor agreement are limited, giving a very small sample. Second, and more importantly, the limited distribution of possessor agreement is very different from the distribution of zero anaphora, hence the tentative inclusion of it in Givón's unstressed/bound pronoun/agreement slot in the scale of phonological size. This actually suggests that we

should remove the possessor agreement category and analyze all instances of possession separately, or combine possessor agreement with zero anaphora since the full NP and pronoun figures include possessors. Another thing to remember is that we are after general patterns that may not fit our expectations in every instance.

So far we have looked only at text A, but we need to consider text B as well, and see how the two compare. In text B the RD values for ellipsis, possessor agreement, pronouns, and full NPs are: 1.15, 1.00, 3.67, and 9.12, respectively, and the PI values are: 1.20, 1.00, 1.27, and 1.92. Table 20 presents the results for text B as actual numbers and percentages:

Table 20. Percentages of High and Low RD and PI values for Text B
(Animate Referents)

	N	RD 1-3	RD > 3	PI = 1	PI = 2
Ellipsis	110	108 98%	2 2%	88 80%	22 20%
Poss. Agr.	15	15 100%	0 0%	15 100%	0 0%
Pronouns	15	13 87%	2 13%	11 73%	4 27%
Full NPs	25	12 48%	13 52%	2 8%	23 92%

Again table 20 suggests that possessor agreement does not belong between ellipsis and pronouns, but all other forms line up nicely, with both ellipsis and possessor agreement more accessible than pronouns, which are more accessible than full noun phrases. All of the figures so far indicate that if possessor agreement belongs in the phonological scale as a separate category, it does not fall between ellipsis and independent pronouns as we would expect to find with subject-verb agreement.

As indicated above, comparing ellipsis with possessor agreement is like comparing apples and oranges. Possession must always be marked with agreement, so ellipsis of a possessor will still leave an overt marking. The ellipsis figures in the above tables represent referents with absolutely no marking whatsoever, so they include no possessors while the possessor agreement figures, on the other hand, include only possessors. Clearly these cannot be compared, especially considering that both the pronoun and full NP figures include possessors. Rather we can examine possessors as a subset of total referents, as in table 21:

Table 21. Percentages for RD and PI of Possessors in Both Texts
(Animate Referents)

Possessor:	N	RD 1-3	RD > 3	PI = 1	PI = 2
Agreement only (ellipted)	9	9 100%	0 0%	8 89%	1 11%
Pronoun (+ agreement)	3	3 100%	0 0%	3 100%	0 0%
Full NP (+agreement)	9	4 44%	5 56%	0 0%	9 100%

These results are what we would expect regarding the relationship between phonological form and accessibility with the exception of the one instance of potential interference among the agreement-only referents. In that particular case, even though there are other semantically compatible characters within the immediate context (PI = 2), the possessor is the most salient. But even if this were not the case, one exception does not destroy an otherwise solid pattern as we have here. We can, however, note that the few pronominal possessors are just as accessible as the ellipted agreement-only examples. I suggest that this is because Adzera pronouns are general enough that they offer very little information that is not already provided by the agreement morphology (cf. section 2.4.3). Full noun phrases on

the other hand are very specific, as indicated by the great statistical difference in accessibility. This begs the question: What is gained by using a pronominal possessor instead of agreement only, and what motivates the choice of one over the other?

These findings suggest that instances in which the possessor is marked only by agreement rightfully belong in the ellipsis category for continuity comparison purposes, and that the fourth category is superfluous. Thus we can collapse tables 19 and 20 into table 22:

Table 22. RD and PI by Text and Reference Form in Three Tiers
(Animate Referents)

		N	RD 1-3	RD > 3	PI = 1	PI = 2
Text A	Ellipsis	44	44 100%	0 0%	36 82%	8 18%
	Pronouns	27	23 85%	4 15%	11 41%	16 59%
	Full NPs	58	36 62%	22 38%	6 10%	52 90%
Text B	Ellipsis	125	123 98%	2 2%	103 82%	22 18%
	Pronouns	15	13 87%	2 13%	11 73%	4 27%
	Full NPs	25	12 48%	13 52%	2 8%	23 92%

Collapsing the tables into three tiers gives results exactly like one would expect on the basis of Givón's scale of phonological size, namely that in both texts ellipsis encodes the most accessible referents, followed by pronouns, and finally full NPs, in terms of both RD and PI.

The two texts have amazingly similar results, except in RD values for full NPs and PI values for pronouns. These minor differences are certainly not problematic. We should expect some differences between texts according to the type of story and the number of characters interacting. Text B has fewer characters than text A, and the main characters

participate in long action sequences before the speaker switches to another character. The speaker of text A, on the other hand, switches from one character to the next much more often, requiring more explicit references and allowing less ellipsis. The "N" column in table 22, above, reflects these facts. Notice the high incidence of ellipsis in text B, compared to text A, in which full noun phrases are more frequent than ellipsed entities. The larger number of characters interacting throughout text A causes full noun phrases to be used in places where RD is low more often than in text B; hence the higher percentage of low RD occurrences of full NPs. We can also postulate that the large number of characters creates more potential interference in cases of pronoun use in text A than in text B. Thus there are logical explanations for some of the differences based on the type of story and the narrator's style, but with or without them, the similarities between the two texts show a very solid pattern. I combine the results from both texts and present them as table 23:

Table 23. Accessibility (RD and PI) for Both Texts
(Animate Referents)

	N	RD 1-3	RD > 3	PI = 1	PI = 2
Ellipsis	169	167 99%	2 1%	139 82%	30 18%
Pronouns	42	36 86%	6 14%	22 52%	20 48%
Full NPs	83	48 58%	35 42%	8 10%	75 90%

5.4.2.2 Accessibility of Inanimate Topics

So far we have examined animate referents in terms of accessibility (RD and PI).

Now we will briefly turn our attention to inanimate referents for comparison purposes. Table 24 presents the accessibility of inanimate topics broken down by form of reference:

Table 24. Accessibility (RD and PI) of Inanimate Topics

	N	RD 1-3	RD > 3	PI = 1	PI = 2
Ellipsis	16	12 75%	4 25%	16 100%	0 0%
Pronouns	6	6 100%	0 0%	6 100%	0 0%
Full NPs	68	19 28%	49 72%	24 35%	44 65%

These results are somewhat surprising in that ellipsed referents appear to be less accessible than pronouns in terms of RD. Note, however, that both types have no potential interference. Why do we have some high RD values with zero anaphora? The ellipsed entities are semantically implied as complements of verbs in the text, so the verbs themselves bring the entities to mind, even if they have not been mentioned in the previous few clauses. The implied entities have been previously evoked and remain salient (or active) in the context by virtue of the circumstances (cf. frame or schema in section 3.4.5), and can be ellipsed in the absence of other potentially interfering referents. For example, the stream of water in the crocodile story (text B) is always salient in the hearer's mind, and a verb indicating descent (into the water) is enough to bring about the image of the crocodile going down into the water without explicitly mentioning the word 'water'. It is precisely this kind of ellipsed reference that often has a higher RD value in the two Adzera texts. In my opinion then, this does not constitute counterevidence to Givón's scale of phonological size.

Givón's method looks at the average accessibility values for each form of reference, rather than examining in detail each reference and the motivation for the reference form. What this method overlooks are the individual instances that do not seem to fit the mold. I just noted examples. Equally noteworthy are the third person pronouns that occur in

environments with RD and PI values of 1 (cf. table 24, above), sometimes even in the middle of series of zero anaphora. This approach cannot explain the motivation of these pronominal appearances. Other types of discourse studies are needed to account for such “anomalies.” They may occur for emphasis, to mark thematic paragraph boundaries, as periodic reminders of the subject, because the narrator is pausing to think about what to say next, or for other reasons. The use of third person pronouns (*araja* ‘he’) in A-12 (see appendix A) appears to mark an episode boundary, and to contrast with A-2.

It is interesting and quite satisfying that both texts show approximately the same distribution of reference forms and roughly similar RD and PI values for each type. A comparison of the two texts in table 25 below suggests that we are seeing a fairly typical distribution of inanimate entities in Adzera narratives:

Table 25. A Comparison of Accessibility of Inanimate Topics in Both Texts

		N	RD 1-3	RD > 3	PI = 1	PI = 2
Text A	Ellipsis	6	5 83%	1 17%	6 100%	0 0%
	Pronouns	3	3 100%	0 0%	3 100%	0 0%
	Full NPs	41	9 22%	32 78%	10 24%	31 76%
Text B	Ellipsis	10	7 70%	3 30%	10 100%	0 0%
	Pronouns	3	3 100%	0 0%	3 100%	0 0%
	Full NPs	27	10 37%	17 63%	14 52%	13 48%

The distribution of forms for inanimate referents is quite different from that which we saw of animate referents. Ellipsis of inanimate topics is much less frequent than ellipsis of animate participants. This is not surprising since we expect animate participants to be more

topical and salient throughout a text than their inanimate counterparts. In fact, the relatively few inanimate referents that are expressed by pronouns and zero anaphora constitute evidence that such is the case. The RD percentages for full noun phrases also bear this out. Animate referents expressed as full NPs are much more accessible on average in terms of RD than inanimate referents expressed as full NPs.

5.4.3 Topic Persistence and Form

5.4.3.1 Persistence of Animate Participants

While it is easy to see the relationship between the two measures of accessibility and the form of reference, it is not as easy to see the relationship between topic persistence and the form of reference. The form that a referent takes seems to be wholly dependent upon identifiability or accessibility, though it is logical that highly persistent topics will remain salient in the text and thus might tend to have low RD values. Using the topic persistence definition from Givón 1983 (no gaps), full NPs have a greater likelihood of having high persistence than less explicit forms, because they tend to occur at the beginning of a string of references while other forms occur in the middle and at the end of a string. On the other hand, full NPs sometimes occur in isolation and have no persistence at all. Using Givón's 1995 definition of TP in which a gap is allowed (i.e. the definition employed in this thesis), it is difficult to make a prediction correlating TP to the form a referent takes. One might make a case for ellipted referents having higher persistence, since they must be established entities and would frequently be main characters. However, just looking at animate entities alone reveals no direct correlation between topic persistence and form of reference, as table 26 shows. Ellipted entities do tend to have high persistence, but not necessarily higher than

Table 26. Persistence of Animate Participants by Text and Form of Reference

		N	TP 0-2	TP > 2
Text A	Ellipsis	44	10 23%	34 77%
	Pronouns	27	7 26%	20 74%
	Full NPs	58	24 41%	34 59%
Text B	Ellipsis	125	18 14%	107 86%
	Pronouns	15	6 40%	9 60%
	Full NPs	25	3 12%	22 88%
Overall		294	68 23%	226 77%

other animate referents. Table 26 shows no predictable overall pattern, even though text A demonstrates a weak pattern with descending values in the right-hand column.

Topic persistence is a measurement of the relative importance of the referent, which I would not necessarily expect to be directly reflected in the form of reference. The most important characters should be more persistent throughout the text than other characters. This seems so intuitive as to be completely obvious, since one way we might judge the importance of a participant is by how often and for how long he shows up in the story. If that is our basis for what constitutes an important character, it would be circular reasoning to claim that this method shows that the most important characters are the most persistent. It is more likely, however, that people have intuitions about which participants are most important, without necessarily knowing exactly why such is the case or how many occurrences a participant has. Therefore, this measurement is a way to quantitatively confirm our intuitions about the relative importance of characters. Table 27 gives both

average TP values and percentage of high TP values (TP > 2)⁵ for main characters vs. supporting characters:

Table 27. Persistence of Main vs. Supporting Characters
(Integrating Percentages and Averages)

		Main Characters			Other Characters		
		N	TP > 2	Avg. TP	N	TP > 2	Avg. TP
Text A	Ellipsis	24	24 100%	8.83	20	10 50%	3.60
	Pronouns	23	19 83%	7.61	4	1 25%	2.25
	Full NPs	21	19 90%	9.62	37	15 41%	2.57
Text B	Ellipsis	110	98 89%	7.90	15	10 67%	4.13
	Pronouns	9	9 100%	8.67	6	0 0%	0.83
	Full NPs	19	18 95%	9.68	6	4 67%	4.50
Overall		206	187 91%	8.35	88	40 45%	3.07

Average TP for main characters is 8.35 clauses, while it is only 3.07 for other characters. To state it another way, main characters have high persistence in 91% of occurrences, while other characters have high persistence in only 45% of occurrences. Here I include in "main characters" not necessarily just one main character, but the two or three main participants in each narrative. In text A that includes the narrator himself (Gadan) and his father (Ayam); in text B it includes Maramai, his wife, and the crocodile. I could have chosen to include Sinkom as a main character of text A in table 27, because of the crucial

⁵ Presenting the data in two different forms, as in table 27, can illuminate slightly different angles on the results, or can confirm the results of one method. The binary method is "blind" to very high TP or RD values, while the average method is not. This blindness can be a good thing (if there are a few unusually high values in a small data sample) or a bad thing (if one sample has consistently much higher "high" values than another), so tabulating the results both ways helps to prevent skewing of the data.

role he played, but decided not to because he was found in only one section, and actual tracked references to him were few. Calculating his persistence separately reveals that he, as a topic, is persistent in 7 out of 9 tracked occurrences, or 78%, and his average TP is 7.78 clauses. This is because of untracked references to him within quotations (sometimes first person references within his own quoted speech). According to table 27, above, he could be included among the main characters, since his TP values fall closer to those of the main characters than those of less important participants. However, rather than making a dichotomy of main vs. other characters, one could treat the importance of participants as a continuum and rank each character's relative importance on the basis of individual TP scores.

Note the close correlation between the average (scalar) method and the high-low percentage (binary) method. It is not absolute, to be sure, but they give very similar results. In the right-hand column in particular, arranging the averages from lowest to highest would also arrange their corresponding percentages from lowest to highest. The methods validate each other and confirm that neither method is giving results that are significantly skewed. Even so, the results, especially in the "main characters" column, show that we cannot take as absolute the ordering that a particular method reveals, nor the ordering found in only one text. If we tried to rank the three types of reference forms according to these persistence results, we would arrive at four different rankings based on text choice and the method of tabulating results. This further confirms my expectations that the form of reference is not directly linked to topic persistence, as defined in this paper. Furthermore, the results given here strongly support the intuition that main characters are much more persistent than supporting characters.

5.4.3.2 Persistence of Inanimate Topics

Topic persistence values for inanimate topics, as with animate referents, show little correlation with the kinds of forms that these referents take. The main benefit of taking a look at inanimate topic persistence values is being able to compare them with the values for animate participants. One would expect animate participants to be more important and, therefore, more persistent, as table 28, in fact, shows:

Table 28. Persistence of Inanimate Topics by Text and Form of Reference

		N	TP 0-2	TP > 2
Text A	Ellipsis	6	3 50%	3 50%
	Pronouns	3	3 100%	0 0%
	Full NPs	41	36 88%	5 12%
Text B	Ellipsis	10	8 80%	2 20%
	Pronouns	3	2 67%	1 33%
	Full NPs	27	18 67%	9 33%
Overall		90	70 78%	20 22%

With an overall average persistence rate of only 22% (the other 78% of inanimate referents decaying immediately), the percentages of high and low persistence values for inanimate referents are opposite to those of animate referents, which have a 77% average persistence rate and only a 23% decay rate. Table 28 also shows no significant correlation between persistence rates and forms of reference.

5.4.3.3 Persistence and the Influence of Animacy

Looking at the persistence rates of animate and inanimate referents separately has revealed no direct correlation between topic persistence and forms of reference. Rather, we have only seen that animate participants are more persistent than inanimate topics, and that the persistence of a participant reflects his relative importance in the discourse. There is, however, an indirect correlation between forms of reference and persistence via animacy. When animate and inanimate referents are counted together, the persistence rates for ellipsis, pronouns, and full NPs are: 79%, 63%, and 46% respectively (although in text B, only 56% of pronouns have high-persistence while 60% of full NPs do). The relation, then, is purely a function of the animate-inanimate statuses of the entities encoded by different forms of reference, rather than an inherent and logical correlation.

5.4.4 Continuity and Grammatical Position

5.4.4.1 Accessibility and Grammatical Position

Now we turn our attention to grammatical role and how it correlates with accessibility. First, we will focus on animate participants. Dividing animate participants into subjects and non-subjects allows us to observe the relative accessibility of discourse referents fulfilling different grammatical roles. We saw in chapter 4 that referents occurring as subjects are overwhelmingly evoked, while new referents in that position are less frequent. Rather, new referents tend to be introduced in grammatical roles other than subject. We should, therefore, expect non-subjects to have a higher percentage of high RD values than subjects, and subjects to have a higher percentage of low RD values than non-subjects.

While this is not the case in every instance, table 29 shows this general pattern:

Table 29. Referential Distance of Animate Subjects vs. Non-subjects

		Subjects			Non-subjects		
		N	RD 1-3	RD > 3	N	RD 1-3	RD > 3
Text A	Ellipsis	40	40 100%	0 0%	4	4 100%	0 0%
	Pronouns	20	17 85%	3 15%	7	6 86%	1 14%
	Full NPs	35	23 66%	12 34%	23	13 57%	10 43%
Text B	Ellipsis	87	87 100%	0 0%	38	36 95%	2 5%
	Pronouns	10	8 80%	2 20%	5	5 100%	0 0%
	Full NPs	14	8 57%	6 43%	11	4 36%	7 64%
Overall		206	183 89%	23 11%	88	68 77%	20 23%

As the reader can see, there is not a highly distinct division between subjects and non-subjects. In many cases the percentage difference is not very significant, and there are two obvious counterexamples to the pattern, namely among the pronouns. In text A, two of the three high RD subjects are the vague third person agents we saw in section 4.3, to which I had to assign the maximum value of 20 because each was the first mention. The explanation for the counterexample in text B is also quite simple: the two high RD subjects are first person references to the narrator himself, who was not part of the story. In each case the first person pronoun is not the subject itself, but embedded within the subject NP as a possessive determiner. I have included these with the subjects, but since the narrator is not a character in the story, it might be better to exclude them entirely.⁶ The distinction between the two grammatical roles is most obvious among full NPs (especially in text B), because full NPs

⁶ To eliminate the dilemma of how to analyze possessors that are participants in the story, it seems legitimate to not analyze any possessors as references, but only count them when figuring the continuity of other references. I might be inclined to do this in future studies, and see how results compare.

can encode referents with either high or low referential distance. Conversely, the distinction tends to be lost among the other two forms of reference because their RD ranges are more restricted. Overall, however, subjects do have a higher percentage of low RD values than non-subjects (89% vs. 77%, respectively). Therefore, these results lend support to Givón's claim that subjects are more continuous than non-subjects in terms of accessibility.⁷

There are very few inanimate subjects in the sample (three, to be exact), so this makes subject and non-subject comparison among inanimate entities difficult and tentative at best. Such comparison does show a lower percentage of low RD values for non-subjects than subjects, tentatively indicating that subjects are more accessible. As we noted above, however, the inanimate values were much lower in accessibility than their animate counterparts.

When animate and inanimate results are combined, some clear patterns emerge. Table 30, below, shows the referential distance of subject vs. non-subject with all discourse entities included. Now we can see clearly that subjects are more accessible (89%) than non-subjects (59%) in terms of RD, just as Givón predicts, but the difference in accessibility is largely tied to the animate-inanimate status of the referents. Again, the most significant differences are found among full NPs, where the range of RD values is broader. While there is a pattern in the distribution of high and low RD values that indicates that subjects (regardless of animate-inanimate status) are generally more accessible than non-subjects in

⁷ Ironically, tabulating PI values shows that animate subjects consistently have more potential interference than non-subjects, though the differences were not significant. Why this is the case is not clear, but I would not expect any correlation between PI and grammatical position, since PI is not a measure of salience but potential for confusion with other referents, which can occur in any position. Therefore, I present only RD and TP values in this section.

Table 30. Referential Distance of Subjects vs. Non-subjects
(Animate and Inanimate Combined)

		Subjects			Non-subjects		
		N	RD 1-3	RD > 3	N	RD 1-3	RD > 3
Text A	Ellipsis	40	40 100%	0 0%	10	9 90%	1 10%
	Pronouns	20	17 85%	3 15%	10	9 90%	1 10%
	Full NPs	37	25 68%	12 32%	62	20 32%	42 68%
Text B	Ellipsis	87	87 100%	0 0%	48	43 90%	5 10%
	Pronouns	10	8 80%	2 20%	8	8 100%	0 0%
	Full NPs	15	8 53%	7 47%	37	14 38%	23 62%
Overall		209	185 89%	24 11%	175	103 59%	72 41%

terms of RD, the difference in accessibility of subjects vs. non-subjects is amplified by the animate-inanimate distinction. Givón himself points out the direct correlation between animacy and grammatical roles in his study of the Ute language (1983:183,5). Table 31 summarizes these patterns:

Table 31. RD Values of Subject vs. Non-subject by Referent Type

	Subjects		Non-subjects	
	N	RD 1-3	N	RD 1-3
Animate	206	89%	88	77%
Inanimate	3	67%	87	40%
All Referents	209	89%	175	59%

In each case subjects tend to be more accessible than non-subjects, confirming Givón's predictions, even though either the difference or the data sample may be small. Also in both grammatical positions animate referents are, on average, considerably more accessible than their inanimate counterparts, as Givón also predicts. Because of this, the paucity of

inanimate subjects and the abundance of inanimate non-subjects magnifies the accessibility difference between subjects and non-subjects. The relationship between animacy and grammatical position is tied very closely with semantic case roles, such as agent or patient. Subjects tend to be agents, which tend to be animate. Inanimate entities tend to be patients or other props, which tend to be grammaticalized as objects or obliques. While other languages may use a passivization strategy to make an inanimate entity the subject of a transitive clause, Adzera does not have this strategy. Therefore, the only place inanimate entities occur as subjects in Adzera is in intransitive clauses, particularly stative ones.

These results correspond well to the assumed familiarity results of chapter 4, which showed that new animate entities more frequently occur in non-subject position and new inanimate entities always occur in non-subject position. The relative inaccessibility of non-subjects compared to subjects, especially inanimate non-subjects, indicates a higher percentage of new referents and a lower percentage of evoked referents fulfilling syntactic roles other than subject.

5.4.4.2 Topic Persistence and Grammatical Position

Givón's predictions pertain primarily to accessibility, but one would also expect subjects to be more persistent than non-subjects, since the most topical entities tend to appear as subjects. As with referential distance one would expect a correlation between animacy and the persistence of subject vs. non-subject. Again, Givón notes this kind of correlation in the Ute language (1983:188-90), and this is indeed what my data show as well. One can factor out the animate-inanimate distinction by looking just at animate entities. Table 32

shows that when inanimate entities are not considered, there is no significant difference between subject and non-subject persistence values:

Table 32. Topic Persistence of Animate Subjects vs. Non-subjects

		Subjects			Non-subjects		
		N	TP 0-2	TP >2	N	TP 0-2	TP >2
Text A	Ellipsis	40	10 25%	30 75%	4	0 0%	4 100%
	Pronouns	20	5 25%	15 75%	7	2 29%	5 71%
	Full NPs	35	12 34%	23 66%	23	11 48%	12 52%
Text B	Ellipsis	87	11 13%	76 87%	38	6 16%	32 84%
	Pronouns	10	5 50%	5 50%	5	1 20%	4 80%
	Full NPs	14	2 14%	12 86%	11	1 9%	10 91%
Overall		206	45 22%	161 78%	88	21 24%	67 76%

The 2% overall difference is not significant and in some subgroups non-subjects show a higher persistence rate than subjects. A comparison of TP values for just inanimate referents actually shows higher persistence of non-subjects than the few subjects in the sample, tentatively indicating that non-subjects are more continuous. But considering all referents together yields the results found in table 33, below. Again we see that the persistence rate is much higher among subjects (77%) than non-subjects (50%), as Givón would predict. Before we saw that there was no statistically significant difference in TP between animate subjects and non-subjects, and that inanimate subjects in this sample (as few as they were) were actually less persistent than inanimate non-subjects. So the subject/non-subject distinction is not significant when the animate-inanimate status of all referents is the same.

But when animate vs. inanimate becomes a factor, the less persistent inanimate referents significantly reduce the persistence of non-subjects overall. Therefore, the persistence of Adzera subjects vs. non-subjects is indeed a function of the distribution of animate and inanimate referents more than anything else. In fact, there is no evidence that factors other than animacy affect the topic persistence of referents fulfilling different grammatical roles.

Table 33. Topic Persistence of Subjects vs. Non-subjects
(Animate and Inanimate Combined)

		Subjects			Non-subjects		
		N	TP 0-2	TP >2	N	TP 0-2	TP >2
Text A	Ellipsis	40	10 25%	30 75%	10	3 30%	7 70%
	Pronouns	20	5 25%	15 75%	10	5 50%	5 50%
	Full NPs	37	14 38%	23 62%	62	45 73%	17 27%
Text B	Ellipsis	87	11 13%	76 87%	48	14 29%	34 71%
	Pronouns	10	5 50%	5 50%	8	3 38%	5 63%
	Full NPs	15	3 20%	12 80%	37	18 49%	19 51%
Overall		209	48 23%	161 77%	175	88 50%	87 50%

5.5 Summary

The quantitative analysis presented in this chapter confirms many of our suspicions and supports many of Givón's claims and predicted orderings. While the non-linear nature of natural discourse presents problems for the clause division, counting methods, and continuity measurements of this kind of approach—which is its major drawback—the averages and percentages still capture the overall tendencies, which are quite useful to study.

Based on the results I make the following general observations about the continuity of referents in the Adzera texts. When ellipted entities are included, animate referents outnumber inanimate referents three to one, and they are more continuous than inanimate referents, both in terms of topic persistence (TP) and referential distance (RD). Ellipted referents are generally more accessible than pronominal referents, and both are more accessible than entities encoded by full noun phrases, which is what Givón's scale of phonological size predicts. However, there appears to be no direct correlation between phonological size and topic persistence, the only correlation at all being a function of animacy. Since persistence measures the relative importance of discourse referents (and not their accessibility in the minds of the hearers), animate entities are more persistent than inanimate entities (as already noted) and main characters are more persistent than supporting characters. Overall, subjects are more persistent than non-subjects, but again this is a function of the distribution of animate and inanimate referents in subject and non-subject positions. Inanimate referents rarely occur as subjects (only 3 out of 90 in these texts), and are less persistent than animate referents, which occur as subjects more often than as non-subjects. Measuring the subject vs. non-subject persistence of just animate or just inanimate referents shows no meaningful pattern. Subjects are also more accessible than non-subjects according to a comparison of RD values, but this is also largely tied to the distribution of animate and inanimate referents. There does, however, appear to be an independent correlation between referential distance and syntactic position, which, I suspect, is related to the introduction of new referents in non-subject positions, as we saw in chapter 4.

CHAPTER 6

FINDINGS AND CONCLUSIONS

6.1 Summary of Findings

Many linguists have written from numerous angles on the subject of given vs. new information, often employing different definitions. This thesis takes the approaches of two of them, Prince and Givón, and applies them to two oral narratives in the Adzera language. Prince is concerned with the assumed status of the referent in the hearer's mind, while Givón is concerned with the topicality or continuity of those referents and how this relates to other factors like form of reference, grammatical position, and animacy.

6.1.1 Assumed Familiarity of Overt Discourse Referents and Their Distribution

Prince's (1981) approach incorporates previous definitions of given and new into an integrated taxonomy of assumed familiarity, with which one can classify overt referents in a discourse. Following this method I assigned each overt referent with one of the following seven classifications: brand-new (BN), brand-new anchored (BN_A), unused (U), inferable (I), containing inferable (I_C), evoked (E), and situationally evoked (E_S). I gave examples of each of these in Adzera, and found them to be similar to the English examples given by Prince, with only a couple of noteworthy exceptions. One was that bare nouns in Adzera, while often indefinite, could be definite also.

The other (and more interesting) exception was with pronouns. When used referentially, pronouns normally refer only to textually or situationally evoked entities—those with clear antecedents. However, there are cases in which the referent is inferable from an indirect antecedent, as in, “I went to the store, but they were out of chocolate chips.” *They* is inferred to be the store owners or managers. In the Adzera data, I found two instances of third person plural pronouns that were not evoked, nor could a clear indirect antecedent be found, so precise identification was impossible based on the utterances in question. In those cases we saw that the identity of the agents encoded by the pronouns was not crucial to the story, and that the author was likely using this construction as the functional equivalent of a passive, which Adzera lacks. Though a case could be made for excluding these occurrences from the analysis, I decided to count them because they denoted specific groups, one of which I believe was later specified. I treated these vague pronouns as inferable entities, because any relevant information about these referents could be inferred.

After classifying the referents I presented the distribution for each class of referent in subject and non-subject positions. I noted that subjects outnumber non-subjects among animate entities 3 to 2. But, overall, overt non-subjects outnumber overt subjects, because of the very few inanimate subjects (8, to be exact), even though overt animate referents outnumber inanimate referents 2 to 1. This is not too surprising when we consider that so many subjects in Adzera discourse are encoded by zero anaphora, and so are not included in the overt entity counts of chapter 4.

Following Prince, I grouped the classifications into three main categories: new (BN, BN_A, U), inferable (I, I_C), and evoked (E, E_S). While new entities constitute a minority

(under 25%) in each grammatical position of all tables, new entities in subject position are much rarer than those in non-subjects positions, supporting the observations of Prince and others that speakers tend to introduce new information in positions other than subjects (see section 6.3, below). The scarcity of new subjects is most apparent among inanimate referents, where all new referents are non-subjects, and the subjects are split evenly between evoked and inferable entities. As noted in chapter 4, this supports Birner's (1994, 1997) observations that inferable entities pattern like evoked entities.

One reason for the scarcity of inanimate subjects is that the subjects of transitive clauses are usually agents and inanimate entities are rarely agents (in fact, never in these texts), so inanimate referents are, for the most part, limited to objects, obliques, and intransitive subjects, as in these texts. Furthermore, Adzera has no passive construction with which to turn a logical object into a grammatical subject. Likewise, Adzera has no dummy subjects or existential or presentational "there" devices (cf. Ward 1999) to move a new logical subject to a non-subject position, thus necessitating the introduction of some new animate referents in subject position (see section 6.3, below).

6.1.2 Continuity of Discourse Referents

Givón (1983) proposes three measurements for determining the continuity of discourse referents: referential distance (RD), potential interference (PI), and topic persistence (TP). The first two measurements relate to accessibility or identifiability of referents in the hearer's mind, while the third is a measure of the importance of the referent in the discourse. I applied these measurements to all "appearances" of discourse entities in the Adzera texts, both overt and implied, with certain exceptions noted in chapter 5. The

inclusion of ellipted references caused the animate to inanimate ratio to increase from 2:1 (overt only) to 3:1, and it also caused subjects to outnumber non-subjects, as we would expect to see.

Givón's (1983) scale of phonological size predicts that zero anaphora will encode more continuous/accessible topics than unstressed/bound pronouns, which, in turn, will encode more continuous/accessible topics than stressed/independent pronouns, which, finally, will encode more continuous/accessible topics than full noun phrases. We saw that to be true in terms of accessibility among the three significant forms of reference in Adzera: zero anaphora, independent pronouns, and full noun phrases. Zero anaphora normally occur when the same referent can be found (even implicitly) in the previous clause ($RD = 1$), and there is no potential interference. They can, however, occur when one or both of these conditions are not met. Pronouns have a bit more freedom, but are generally found within a few clauses, and have low interference. While first and second person pronouns are common, third person pronouns are not as common, zero anaphora being the preferred form. Full noun phrases can occur anywhere regardless of referential distance, but usually must occur when either RD or PI is high, and frequent occurrence when both are low would be very odd.

The correlation between these forms of reference and topic persistence is not as obvious, and was clearly related to the animacy of the referents, since animate referents tend to be more important topics than inanimate entities, and, therefore, more persistent. Because animate entities are ellipted much more often than inanimate entities, which are typically encoded by full noun phrases, entities encoded by zero anaphora did ultimately prove to be

more persistent, on average, than pronouns, and finally full NPs. However, there is no direct correlation between topic persistence and form of reference. We did see that topic persistence is useful for ranking the relative importance of discourse participants.

Givón (1983) also predicts that subjects will be more continuous/accessible than objects, which will be more continuous/accessible than other sentence constituents. He also makes related predictions regarding agents, recipients, patients, etc., and animate vs. inanimate entities. Though I found exceptions to this, particularly comparing subjects and non-subjects among animate pronouns, overall the data confirms that subjects are more accessible than non-subjects in terms of RD, and the exceptions, which include the vague third person plural pronouns discussed above, are easy to explain. This corroborates the results summarized in 6.1.1, since new referents are less accessible in terms of RD than most given referents, and are usually presented in non-subject positions. The differences in RD between subjects and non-subjects are most dramatic when animate and inanimate entities are combined, allowing animate-inanimate distribution to exert its influence. Topic persistence of subjects vs. non-subjects shows no significant difference in these two Adzera texts as long as animate status is held constant. But when the distribution of animate vs. inanimate referents become a factor, subjects (almost entirely animate referents) prove to be considerably more persistent than non-subjects, confirming Givón's prediction but on the basis of animacy alone.

6.2 Conclusions and Applications

Applying these two separate but related approaches to Adzera text proves to be useful. Although Givón's linear approach to hierarchical discourse presents some problems

regarding the counting of referents, the method still captures the general patterns of continuity among discourse entities in Adzera. This study confirms Givón's (1983) predictions, but demonstrates that some relationships, such as that between topic persistence and grammatical position or topic persistence and form of reference, have no obviously inherent correlation, but appear to be simply the results of other factors, in these cases, animacy. It further confirms that the correlation between accessibility and grammatical position is also influenced by animacy. Both methods highlight the tendency for new referents to appear in non-subject positions (at least in SVO languages like Adzera). Prince's (1981, 1992) approach helps the analyst to thoroughly consider the given-new status of referents, and to look for possible antecedents, direct or indirect. This classification method is sensitive to inference, something which Givón's quantitative approach must necessarily ignore. Neither method would predict pronominal forms being used for first mention, as they normally must have an antecedent (an indirect one at the very least). However, Adzera speakers can apparently use such agent pronouns in a special, attributive way, causing the construction to function like a passive, which the language lacks. These pronouns do not actually introduce new information, as the agents are meant to be largely ignored. These unexpected exceptions do not present a major problem for either method, however. Because the two approaches are designed to examine different aspects of information structure and different subsets of the data, I see them as complementary. They enable the analyst to approach the data from different perspectives to achieve an overall picture.

Knowing what forms occur in which circumstances is useful for translating materials naturally into Adzera. We saw that third person pronoun use is relatively infrequent

compared to zero anaphora for encoding the most identifiable topics and participants in this genre. Translating source language pronouns into Adzera as pronouns in many cases would sound stilted to native speakers, since, for third person, ellipsis is the preferred encoding strategy. Something as simple as this can make a translation sound either completely foreign or quite natural, depending on whether the translator is sensitive to the differences in encoding strategies between the two languages. Therefore, this kind of study can be very important to Adzera speakers and translators.

6.3 Future Research Directions

A study of this kind can cover many different aspects, and this present study has just scratched the surface. In this beginning I have only dealt with a small sample of oral narratives. The corpus of narrative texts could be expanded to see if the same findings hold true in other narratives. Future research could focus on other genres as well, such as procedural texts, different kinds of hortative texts like sermons or persuasive speeches, expository texts describing aspects of Adzera culture, or a comparison of different genres.

With regard to Givón's predictions, I have focused on the relationships between topic continuity and form of reference and topic continuity and subjects vs. non-subjects. It should prove insightful to examine subjects, objects, and obliques, and how they relate to continuity, though a larger data corpus would give better results. Furthermore, I feel that the relationship between animacy, semantic role (agent, recipient, patient, etc.), and grammatical position deserves a closer look, again when more data is examined, allowing finer distinctions to be made without making each sample too small. Dividing noun phrases into the categories of bare nouns, proper nouns, and definite and indefinite NPs could prove

insightful, as well as looking at sentence position and non-canonical word orders (i.e. the fronting of certain constituents).

One study that I believe may prove to be particularly insightful, in the approaches of both Prince and Givón, is examining a three-way¹ split between agents (subjects) of transitive clauses (A), subjects of intransitive clauses (S), and objects of transitive clauses (O). Du Bois (1987) shows that new information tends to appear as objects and intransitive subjects rather than as agents of transitive clauses. A tentative examination of the six new referents introduced as subjects in the two texts (tables 9 and 10) reveals that four of them are intransitive subjects and the other two are “agents” of the verbs ‘hear’ and ‘see’. This proves interesting because these verbs have fewer transitive properties than many other verbs (see Hopper and Thompson 1980 for a discussion of the scalar properties of transitivity). Much more data would be needed to corroborate this kind of trend in Adzera, but it certainly is a promising area for future research. An integration of the approaches of Givón and Prince also may be more helpful than approaching different subsets of the same data from different perspectives. One such approach could primarily follow Givón’s method while using Prince’s taxonomy to classify the same referents. Zero anaphora could be dealt with easily and would be classified as evoked.

As the reader can see, there are many directions that could be taken in future research in regard to information structure and topic or participant tracking and distribution. The broader field of discourse study is also wide open for future study in regard to the Adzera language. There are a number of approaches that could be applied to Adzera texts of various

¹ At least three-way (A, S, O), but possibly including other distinctions mentioned, such as recipients and obliques.

genres. Finally, in addition to the discourse level, there are areas of syntax that have not been explored in much detail, such as relative clauses and complement clauses, as noted in chapter

2. These are just some of the topics that are worthy of more attention. Indeed, the avenues for further research seem endless, and I hope to pursue some of them in future studies.

APPENDIX A

TEXT A: OPPOSITION TO MISSION WORK

Told by Gadan Ayam, January 2000, Binimamp village.

1. *Orait, ramaŋ' miamun da iba mpai, impa wa gamp.*
 alright father-1pss before TIME R-come stay R-stay CMPLT village
 'All right. Many years ago my father came and stayed in the village.'
2. *Da wa', anuŋ' ifa i gum mision u.*
 and (um) NEG R-go for work mission NEG
 '(He) did not go to do mission work.'
3. *Itaŋin gum mision rai, da iba ani*
 R-leave work mission away and R-come here

i mpada gan arun gamp wasa'.
 PURP stay-NMZ 3pss way-3pss village inside
 '(He) left mission work and came here to stay for good in the village.'
4. *Orait, iru' ya irim dzaf a Guaruk.*
 alright R-go.down and R-put fire (uh) Guaruk
 'All right. (He) went down and burned (*kunai* grass) at Guaruk.'
5. *Da wa', dzar' iwa'.*
 and (um) pig R-come.out
 'and a pig came out.'
6. *Ini ntaŋa dzar' da irut raingan Augustin.*
 R-want spear-NMZ pig and.so R-accompany brother-3pss Augustin
 '(He) wanted to spear the pig so (he) went with his brother Augustin.'
7. *Orait, Augustin imunti da iri ntaŋa dzar'*
 alright Augustin R-stand and R-want spear-NMZ pig

da ibutip intaŋ ramaŋ' i taramapruf.
 but R-instead R-spear father-1pss with metal.spear
 'All right. Augustin stood there and tried to spear the pig but instead (he) hit my father with the spear.'

8. *Da ramaŋ' ... ramaŋ' itapu waia' rai*
and father-1pss father-1pss R-throw metal.tip away

da ifa hakim a, Augustin banjin.
and R-go hook (uh) Augustin hand-3pss

'And my father threw the metal spear point away and (it) went and caught Augustin's hand.'

9. *Orait, ramaŋ' igiŋ' puatsi isaŋ' a,*
alright father-1pss R-lie kunai.grass R-enough (uh)

raingan Augustin ntaŋa i bampun sib imiŋ' Guaruk.
brother-3pss Augustin R-sew Obj. abdomen-3pss CMPLT R-at Guaruk

'All right. My father laid in the *kunai* grass until his brother Augustin had sewn up his abdomen at Guaruk.'

10. *Intsup sib, bampun arigi iri', ramaŋ' iba.*
R-finish CMPLT abdomen-3pss that.one R-heal father-1pss R-come

'When (he) had finished, the abdomen healed and my father came here.'

11. *Iba mpa gamp wasa' Binimamp.*
R-come stay village inside Binimamp

'(He) came and stayed in Binimamp.'

12. *Araŋa ufi' rusakan da araŋa uri' ifa gum mision imiŋ' a...*
3sg carry rucksack-3pss and 3sg get.up R-go work mission R-at (uh)

'He carried his rucksack and he got up and went to do mission work.'

13. *Ini fada Dzigintsuan*
R-want go-NMZ Dzigintsuan

da ifa sau garam Dzigintsuan gampan.
and.so R-go look.for men Dzigintsuan village-3pss

'(He) wanted to go to Dzigintsuan so (he) went to find Dzigintsuan village.'

14. *Da garam Dzigintsuan ini bina',*
and men Dzigintsuan R-say like.this

'But the men of Dzigintsuan said,'

15. <<Aaa, u ba ibinigi da Buaruŋ' buŋ' aba
 ah 2sg (R)come R-like-that and Buarung FUT IR-come
 da warib is aga fafarab.
 and continue.as.always kill 1pl.ex in.pieces
 '«Ah, you've come and so the men of Buarung will surely come and massacre us as
 always.»'
16. Watip wafan.>>
 IMP-repeat IMP-go
 '«Go back.»'
17. Da ramaŋ' ibugin a... ramaŋ' itsaja
 and father-1pss R-not.want (uh) father-1pss R-see
 rib Dzigintsuan buginda wa', ramaŋ' i
 3pl Dzigintsuan not.want-NMZ (um) father-1pss INTENS
 da itip iba ya iba yab ifa Sukurum.
 and R-repeat R-come and R-come go.up R-go Sukurum
 'My father saw that the people of Dzigintsuan didn't want him (to stay) so (he) came
 back and went up to Sukurum.'
18. Ifa mpa... irut ŋawai impa Sukurum.
 R-go (stay) R-accompany Ngawai R-stay Sukurum
 '(He) went and stayed with Ngawai in Sukurum.'
19. Raiyi da wa', ribigi ini nan
 later TIME (um) 3pl R-say talk
 da itip iyu ramaŋ' iba fawa' Saum.
 and R-repeat R-get father-1pss R-come go-arrive Saum
 'Later they spoke and brought my father back to Saum.'
20. Da impa Saum imiŋ' ŋunmanu'.
 and R-stay Saum R-at Ngunmanu
 '(He) stayed in Saum at Ngunmanu.'

21. *Garam Saum igi impa rugan gampan*
 men Saum there R-stay different village-3pss

ibi rib Siruwarang? gin aga
 R-like 3pl Siruwarang 3obj. over.there

da ramar? runta? impa wap wasa? ibiani.
 and father-1pss alone R-stay forest inside R-like-here

'The people of Saum lived in a different village like over in Siruwarang, for example, and my father lived alone in the woods like over here.'

22. *OK, ramar? iba ya iba yu dzi imin? gamp ani.*
 OK father-1pss R-come and R-come get 1sg R-from village here

'OK. My father came and took me from here.'

23. *Da dzi irut iyab.*
 and 1sg R-accompany R-go.up

'And I went up with (him).'

24. *Iyu a... Dzi binanga? a? Gadan.*
 (R-get) (uh) 1sg name-1pss (uh) Gadan

'(He) took... My name is Gadan.'

25. *Da wa? dzi irut ramar? iyab.*
 and (um) 1sg R-accompany father-1pss R-go.up

'I went up with my father.'

26. *Aga fa mpai.*
 1pl.ex go stay

'We went and stayed.'

27. *Dzi impa tayar? ramar? da ramar? inaj aga gumar?.*
 1sg R-sit look.after father-1pss and father-1pss R-do 1pl.ex work-1pss

'I took care of my father and he did our work.'

28. *Rinaŋ' impa gamp ani,*
 mother-1pss R-stay village here
da wa' dzi da ramaŋ' aŋu, aga mpa waguŋ' aga.
 and (um) 1sg and father-1pss only 1pl.ex stay above there(far)
 'My mother stayed here, and just my father and I, we stayed up there.'
29. *Aga mpai ifa fa, da Siŋkom, garam tsira' maŋan biŋaŋgan,*
 1pl.ex R-stay R-go go and Sinkom man big a/one.of name-3pss
isu ruruai da ini bina',
 R-become officer and R-say like.this
 'After we had stayed there for a while, Sinkom—(that's) the name of one of the big men (i.e. leaders)—became the local colonial government officer, and said,'
30. <<O, *agam mision ruas igi!*
 oh 2pl missionary and.company there
Agam iba i suda umpur i agai.
 2pl R-come PURP become-NMZ untrue PREP 1pl.ex
 '<<Oh, you missionaries! You came to deceive us.'
31. <<Da *aga wa', naŋga' gadan da apugaŋ',*
 and 1pl.ex (um) thing-1pss eat-NMZ and meat-1pss
aga westim natiŋ aŋu i agam.
 1pl.ex waste for.nothing only PREP 2pl
 '<<And our food and meat, we are just wasting on you.'
32. <<Arani *da dzi buŋ' ais agu ya dzi naga agu*
 today TIME 1sg FUT IR-kill 2sg and 1sg FUT-eat 2sg
da dzi nayu gudzum da dzi natsara ayab uŋar gudzum.
 and 1sg FUT-take head-2pss and 1sg FUT-offer IR-go.up house head-3pss
 '<<Today I will kill you and eat you, and I will take your head and offer (it) up on the roof of the house.'
33. <<Da *wa', ibiŋimp aba da naba aga u gudzum arigi.>>*
 and (um) eagle IR-come and FUT-come IR-eat 2sg head-2pss right.there
 '<<And the eagles will come and eat your head there.>>'

34. *Da raingan Rampias itip ini ibiani, irut Gurump,*
 and brother-3pss Rampias R-again R-say R-like-this R-with Gurump
ini bina?
 R-say like.this
 'Then his brother Rampias, together with Gurump, replied like this.'
35. <<*Wafa ais mision ya waba agan.*
 IMP-go IR-kill missionary and IMP-come IR-eat
 '<<Go and kill the missionary and come and eat (him).>>'
36. <<*Nantsup sib, orait aga...>>*
 HORT-finish CMPLT alright 1pl.ex
 '<<When (you) are finished, all right, we...>>'
37. -- *raingan stret -- ini bina?*
 brother-3pss true R-say like.this
 '-- his actual brother, (mind you) -- he said.'
38. <<*Wais mision ya wagan.*
 IMP-kill missionary and IMP-eat
 '<<Kill the missionary and eat (him).>>'
39. <<*Da dzi burj' atip ais agu ya dzi naga agu*
 and 1sg FUT IR-again IR-kill 2sg and 1sg FUT-eat 2sg
nasarj' mision nangan muja?>>
 FUT-enough missionary thing-3pss retribution
 '<<Then I will kill you in return and eat you as payment for what you have done to the missionary.>>'
40. *Da wa?, Sinkom wafa ais agai u.*
 and (um) Sinkom NEG-go IR-kill 1p.ex NEG
 'So Sinkom didn't kill us.'

41. *Sinkom impa, i biloŋ wanem ibi Rampias awa' Gurump*
 Sinkom R-sit CAUSE because(Pidgin) R-be.like Rampias and(um) Gurump
irim a ramaŋ' sib.
 R-help (uh) father-1pss PRT

'Sinkom didn't do anything because Rampias and Gurump protected my father.'

42. *Da ramaŋ' ini da dzi,*
 and father-1pss R-say to 1sg

'Then my father said to me,'

43. <<*Bida agi tsananda simpup a num badan,*
 be.like-NMZ 1pl.in see-NMZ drum and conk.shell come-NMZ
da warunt ya wafan ya iya' warut ŋawai wampa.
 then IMP-run and IMP-go and (um) IMP-with Ngawai IMP-stay

'<<If we hear the drums and conch shell coming, then run away and stay with Ngawai.'

44. <<*Da u saŋ' u mudan mpui Leron??>>*
 but 2s (R)can 2s cross-NMZ river Leron

'<<But can you cross the Leron River?>>'

45. *Da dzi ini,* 46. <<*Dzi isaŋ'.*
 and 1sg R-say 1sg R-can

'And I said, <<I can.'

47. <<*Dzi buŋ' amap mpui da dzi nawa' Sukurum.>>*
 1sg FUT IR-swim river and 1sg FUT-arrive Sukurum

'<<I will swim the river and reach Sukurum.>>'

48. *Da Mugusa irunt iba.*
 and Mugusa R-run R-come

'Then Mugusa came running.'

49. *Gurump narungan bijangan Mugusa.*
 Gurump child-3pss name-3pss Mugusa

'Gurump's son's name was Mugusa.'

50. *Da isalim arigi, ini bina?*
and R-send that.one R-say like.this
'(He) sent him, saying,'
51. <<*Warunt da wani da mision i ya'*
IMP-run and IMP-say to missionary Comp (um)
Sijkom burj' a'... nabia' nais a', mision da naga
Sinkom (FUT) (IR-) FUT-come FUT-kill (uh) missionary and FUT-eat
da gubu' arani.
TIME day this.one
'<<Run and tell the missionary that Sinkom will come and kill him and eat (him) today.'
52. <<*Da natip rungan ma nafan ya nagup.>>*
and.so HORT-prepare self-3pss or HORT-go and HORT-hide
'<<So (he) should get himself ready or go and hide.>>'
53. *Da ramaj' ini da dzi,*
and father-1poss R-say to 1sg
'And my father said to me,'
54. <<*Dzi wasaj' runtan u.*
1sg NEG-can run-NMZ NEG
'<<I won't run.'
55. <<*Jisu guman, dzi burj' amamp.* 56. *Da agu?*
Jesus work-3pss 1sg FUT IR-die but 2sg
'<<It's Jesus' work; I will die. But you?'
57. <<*U burj' arunt i ya' wa'a da gayangam ŋawai*
2sg FUT IR-run PURP (um) arrive-NMZ to uncle-2pss Ngawai
ya Sukurum.>>
LOC? Sukurum
'<<You will run to reach your uncle in Sukurum.>>'

67. *Sunta' maŋa fa sib.*
 week a go CMPLT
 'One week passed.'
68. *Raiyi Sip ipinis a'... Sip iriŋaŋ' i aga mpada puatsi*
 later Sip R-finish (uh) Sip R-hear Comp 1pl.ex stay-NMZ kunai.grass
da Sip impa wa' ABCO.
 and Sip R-stay (um) ABCO (company name)
 'Later (Sip finished...) Sip heard that we were living in the bush and he was at ABCO.'
69. *Da araŋa ba*
 and 3sg come
da iba asu ramaŋ' bugisan imiŋ' a' ŋunmanu'
 and R-come carry father-1pss box-3pss R-from (uh) Ngunmanu
 'So he came and carried my father's things from Ngunmanu.'
70. *Da aga fa Gimigim.*
 and 1pl.ex go Gimigim
 'And we went to Gimigim.'
71. *Aga fa mpa Gimigim, rib a', Wagiasa' gampan.*
 1pl.ex go stay Gimigim 3pl (uh) Wagiasa village-3pss
 'We went and stayed in Gimigim, a village of the Wagiasa people.'
72. *Aga mpa arigi, orait,*
 1pl.ex stay right.there alright
ribigi ya'... irurub a', rib Saum.
 3pl (um) R-baptize (uh) 3pl Saum
 'We stayed there, and they baptized the people of Saum.'
73. *Irurub santan aru.*
 R-baptize all completely
 '(They) baptized everyone.'

74. *Irurub* *i* *Saum,* *Sukurum,* *Fayanj,* *Ŋariaway?*
 R-baptize Obj Saum Sukurum Fayang Ngariawang
- Sukurum,* *Sia?sia?* *nam...* *Papuai,* *Sisuk,* *Mugaba.*
 Sukurum Sirasira what... Papuai Sisuk Mugaba

‘(They) baptized Saum, Sukurum, Fayang, Ngariawang, Sukurum, Sirasira, who else... Papuai, Sisuk, Mugaba.’

75. *Santan* *aju*¹ *da* *irurub* *imij?* *a?* *rib* *Saum* *gampan.*
 all completely and R-baptize R-at (uh) 3pl Saum village-3pss

‘All of them (they) baptized at Saum village.’

76. *Ramaj?* *ruas* *ani* *ifan* *da* *irurub.*
 father-1pss pl here R-go and R-baptize

‘Our fathers here went and baptized (them).’

77. *Da* *tuasaraŋan*² *da* *Sijkom* *ini* *nangan*
 and 3pl and Sinkom R-say talk-3pss
- aru* *a?* *isa,* *tsarifa* *ruan* *i* *isa* *ramaj?* *igi,*
 SPEC (uh) kill-NMZ exalt-NMZ self-3pss about kill-NMZ father-1pss there
- da* *nan* *iwa?* *gin.*
 and talk R-arrive 3obj

‘Sinkom made his boastful threats about killing my father and they heard about it.’

¹ Or possibly: ...*Papuai, Sisuk, Mugaba*—*santanaju*—*da irurub imij? a? rib Saum gampan*. ‘...*Papuai, Sisuk, Mugaba*—all of them—and they baptized them at Saum village.’ However, the intonation and pauses indicate the clause structure given in the text above, since there are pauses between *Papuai, Sisuk, Mugaba* and *santanaju*, but there is not even a slight pause between *santanaju* ‘all of them’ and the conjunction *da*.

² It is not clear to me whether *Da tuasaraŋan* ‘and they’ is a false start (‘And they- And Sinkom made...’), the start of the next sentence plus the late insertion of an overt object for the previous clause as an afterthought (‘And- them. And Sinkom...’), an instance of contrastive focus involving those who went to baptize (‘And them, Sinkom made his boastful threats...and they heard about it.’), or whether the narrator is including the people of Saum with Sinkom (‘They and Sinkom made their boastful threats...’), though this does not fit the story as told. The intonation and volume of *Da tuasaraŋan* ‘and they’ indicate the beginning of a new clause, and there is no pause between *Da tuasaraŋan* and *da Sijkom*. It appears that *Da tuasaraŋan* ‘and they’ is a mistake and that it is not relevant to the text.

78. *Nan ani³ ibawa' arigi.*
 talk this R-come-arrive that.there

'That's the end of the story.'

79. *Teŋkyu.* 80. *Ayam stori-gan igi.*
 thank.you Ayam story-3pss there

'Thank you. That's Ayam's story.'

³ Or *Na nan* 'and talk' (Pidgin: *na* 'and'), giving *Na nan ibawa' arigi* 'That's the end of the story' (same meaning).

APPENDIX B

TEXT B: MARAMAI SPEARS A CROCODILE

Told by Wari Muruk, October 2001, Binimamp village.

1. *Dzi biyanga? Wari.*
1sg name-1pss Wari
'My name is Wari.'

2. *Dzi ni fisa i nan dzi buŋga? Maramai'*
1sg want tell-NMZ Obj story 1sg inlaw-1pss Maramai
mpada miŋ'a wa' Riara.
stay-NMZ at-NMZ (um) Riara
'I want to tell you a story about my father-in-law Maramai who lived in Riara.'

3. *Da inaj gum Fraide sib da irut finingan ifa*
and R-do work Friday CMPLT and R-accompany wife-3pss R-go
i sauda gan yafasan i gadan.
PURP look.for-NMZ 3pss fish-3pss PURP eat-NMZ
'(He) finished the Friday community work and went with his wife to look for fish to eat.'

4. *Ifan da itsaja mpu maran manjan da irunt in iru'.*
R-go and R-see water spring a and R-ran 3obj R-down
'(He) went and saw a spring and followed it down.'

5. *Ifan da itsaja fugai manjan mpada mpui maran arigi wasa'-gan.*
R-go and R-see crocodile a sit-NMZ water spring that middle-3pss
'(He) went and saw a crocodile sitting in the middle of that stream.'

6. *Da inaj ruŋum tsira', ipanu ruŋum tsira' ifa fa wa'*
and R-make deep.spot very R-dig deep.spot very R-go go arrive
da igiŋ' gin da iwanjun.
and R-lay 3obj and R-pregnant
'(She) made it very deep; (she) dug and dug it until there was a very deep hole, and (she) lay in it and was pregnant.'

7. *Iwarjun sib da ini apir'a narun*
 R-pregnant CMPLT and R-want give.birth-NMZ baby-3pss
da iyab ifan ifari git.
 and R-go.up R-go R-build nest
 '(She) was already pregnant and (she) wanted to lay her eggs, so she went up and built a nest.'
8. *Da iyu git igi da yari*
 and R-took nest that and R-build
ibi ya' unjar garam gir'a gin ugu.
 R-like (uh) house people sleep-NMZ 3obj before
 '(She) took the nest and built it like the houses that people sleep in.'
9. *Iyu git igi iyab ya yari ya irua' iru' ifan*
 R-took nest that R-up and R-build and R-fall R-go.down R-go
ya itip ifa yu a' puatsi iba ya iba yab
 and R-repeat R-go get (uh) grass R-come and R-come up
ya irim ya irua' iru' ba.
 and R-put and R-fall R-go.down come
 '(She) took the nest up and built it, and (she) went down and brought back some more grass and came back up and put (it on the nest) and came back down.'
10. *Ibinigi gin da wa' dzi bunga' Maramai' iba*
 R-like-that 3obj and (uh) 1sg inlaw-1pss Maramai R-come
da iba tsarja wa' git igi da ini gan a' dzar' gitan gin.
 and R-come see (uh) nest that and R-say 3pss (uh) pig nest-3pss 3obj
 'That's what was happening and my father-in-law Maramai came along, and (he) came and saw the nest, and thought it was the nest of a pig.'
11. *Iba ba da iba tsarja fugai fagan.*
 R-come come and R-come see crocodile foot(print)
 '(He) kept coming and then (he) saw crocodile tracks.'

12. *Da ini bina?*, <<A'a fugai ani.>>
and R-say like.this NEG crocodile here
'And (he) said, <<No, this is a crocodile.>>'
13. *Da ini da finingan,*
and R-say to wife-3pss
<<Watip wafan da wafa munti gunti' aga da dzi nafa tsajan.>>
IMP-repeat IMP-go and IMP-go stand far over.there and 1sg FUT-go see
'And (he) said to his wife, <<Go back and stand way over there and I will go and look.>>'
14. *Da arajan itip aru iba.*
and 3sg R-careful only R-come
'He came very quietly.'
15. *Itip iba, iba tsajan.*
R-careful R-come R-come see
'(He) crept closer and looked.'
16. *Da fugai irijant i riririj badan*
and crocodile R-hear Obj/Comp noise come-NMZ
da araja irua' iru' mpui ya igij' mpui.
and 3sg R-fall R-down water and R-lay water
'The crocodile hear a/the noise coming and she went down into the water and lay there.'
17. *Igij' mpui igi aru ntras.*
R-lay water there only still
'(She) lay absolutely still in the water.'
18. *Maramai' iba tsaja sib ya Maramai' ifa da ifa tip ini*
Maramai R-come see CMPLT and Maramai R-go and R-go repeat R-say
da finingan, <<Wamunti da dzi nafa tsaja fugai igij' mpui.>>
to wife-3pss IMP-stand and 1sg FUT-go see crocodile R-lay water
'When Maramai had come and seen it, he went back and told his wife, <<Stand here and I will go and see the crocodile sleeping in the water.>>'

19. *Araña tip iba da iba munti da itsaṅa fugai rugij'an.*
 3sg repeat R-come and R-come stand and R-see crocodile CONT-lay-NMZ
 'He came back and stood there and saw the crocodile still sleeping.'
20. *Da araṅa ba munti da ifinti siṅan.*
 and 3sg come stand and R-charm spear
 'He came and stood and put a ritual blessing on the spear.'
21. *Ifinti siṅan igi sib da imunti da intaṅ'.*
 R-charm spear that CMPLT and R-stand and R-spear
 'When (he) had blessed the spear, (he) stood and speared (the crocodile).'
22. *Ini ntaṅ'a fugai da intaṅ' uprats.*
 R-want spear-NMZ crocodile and R-spear completely.through
 '(He) wanted to spear the crocodile, and (he) speared (it) all the way through.'
23. *Ya fugai ruṅ' isu dubidubi da irunt ifan.*
 and crocodile CONT R-be flailing and R-run R-go
 'The crocodile was still flailing about and (he) ran away (from it).'
24. *Da ifa munti da fugai iyai'.*
 and R-go stand and crocodile R-cry.out
 '(He) went and stood (away from it), and the crocodile bellowed.'
25. *Iyai' nifun tsira'.*
 R-cry.out mouth-3pss big
 '(She) bellowed really loudly.'
26. *Ifa ifa ya araṅa yitam i bampu sap iyab.*
 R-go R-go and 3sg R-turn Obj belly white R-go.up
 'After a while she turned belly up.'
27. *Ya imamp funub.*
 and R-die dead
 'And (she) died.'

28. *Da Maramai' itip iba.*
and Maramai R-repeat R-come
'Then Maramai came back.'
29. *Ya iba tsajan da itus apruf sinuŋ'.*
and R-come see and R-pull spear out
'(He) came and looked and pulled out the spear.'
30. *Da wa' inu' finiŋan, <<Waba. Dzi intaŋ' fugai ugu furub wa.*
and (uh) R-call wife-3pss IMP-come 1sg R-spear crocodile that(prev) dead CMPLT
Wabai. Agi dzuŋa gin yaban.>>
IMP-come 1pl.in pull-NMZ 3obj go.up-NMZ
'Then (he) called to his wife, <<Come. I already speared the crocodile to death. Come. Let's pull it up.>>'
31. *Da ibri' yuntap in iyab.*
and R-together R-pull 3obj R-go.up
'So together (they) pulled it up.'
32. *Iyuntap in iyab ifa ibinigi*
R-pull 3obj R-go.up R-go R-like-that
da ini da fin- iguti finiŋan, <<U saŋ' u tariŋ'an?>>
and R-say to (unfinished) R-ask wife-3pss 2sg able 2sg carry-NMZ
'When (they) had pulled it up, (he) asked his wife, <<Can you carry it on your back?>>'
33. *Da finiŋan ini bina', <<Mai', fugai igi tsira' biŋan igi ya!*
and wife-3pss R-say like.this Mai crocodile that big very that EXCL
Dzi wasaŋ' i dzi bantiŋ'an u! Buŋ' abarabin da dzi.>>
1sg NEG-able Comp 1sg carry-NMZ NEG FUT IR-heavy for 1sg
'And his wife said, <<Mai, that crocodile is huge! I can't carry it. (It) will be too heavy for me.>>'

34. *Da ini bina?*, <<A'a, u' buy' abantiq' ayab
and R-say like.this NEG 2sg FUT IR-carry IR-go.up
da dzi buy' anjif i bampan tsira' aruani nayab
and 1sg FUT IR-carry Obj tail-3pss big this.here FUT-go.up
da agi nampru' ruarj' yurja fadan.>>
and 1pl.in FUT-with RECIP-1pss walk-NMZ go-NMZ
'But (he) said, <<No, you carry (it) on your back, and I will carry its big tail here, and we will walk along together.>>'
35. *Da ini bina?*, <<Ari', wara' gin wayab da dzi nafits atsanjan.>>
and R-say like.this OK IMP-lift 3obj IMP-go.up and 1sg FUT-carry IR-see
'So (she) said, <<OK, lift it up and I will see if I can carry (it).>>'
36. *Ifits iyab ibinigi da gabungan, arajan Maramai?, iyu bampan igi*
R-carry R-go.up R-like-that and husband-3pss 3sg Maramai R-take tail-3pss that
da irut iyurj ifa fa fa-a, ifawa' magamarj'
and R-with R-walk R-go go go-o R-go-arrive dry.riverbed
da ifa gu ruan mpada asap.
and R-go with RECIP-3pss sit-NMZ rest
'(She) carried (it) like that, and her husband, Maramai, carried the tail and (he) walked with (her) a long way until (they) came to a dry riverbed, and (they) rested together.'
37. *Ifa mpa asapan sib da itip yuri? ya itip ifits.*
R-go sit rest-pss CMPLT and R-repeat R-get.up and R-repeat R-carry
'After (they) had rested, (they) got up again and (she) continued to carry (it).'
38. *Da gabun itip yapij' bampan arigi.*
and husband-3pss R-repeat R-carry tail that.one
'And her husband continued to carry the tail.'
39. *Ya irut iyu ifa-a ini wa'a gamp tsira' a'*
and R-with R-take R-go R-want arrive-NMZ village big (uh)
wa'a fada gamp.
arrive-NMZ go-NMZ village
'Together (they) took it along until (they) approached the village.'

40. *Da rīb gamp igi itsaṅa ba.*
and 3pl village there R-see come
'The people of the village saw (them) coming.'
41. *Riara igi itsaṅa ba da iyai' gin, <<Ayoiee! Maramai' irut finin*
Riara there R-see come and R-cry.out 3obj EXCL Maramai R-with wife-3pss
ifa ntaṅ' fugai ya itariṅ' iba, ifits iba.>>
R-go spear crocodile and R-carry R-come R-carry R-come
'(The people of) Riara saw (them) coming and cried out about it, <<Ayoiee! Maramai went with his wife and speared a crocodile, and they carried (it) back.>>'
42. *Ya iyu iba ya iba taṅin iru' da impa asapan.*
and R-take R-come and R-come leave R-go.down and R-sit rest-3pss
'(They) brought (it) and put (it) down and sat down to rest.'
43. *Da numuntugan Kakarak da wa' Ati' itsaṅa da ini bina',*
and brother.in.law-pss Kakarak and (uh) Ati R-see and R-say like.this
<<Ake, agi numuntuga' araga ifa ntaṅ' agi apu-ga'.>>
EXCL 1pl.in brother.in.law-1pss that.there R-go spear 1pl.in animal-1pss
'His brothers-in-law Kakarak and Ati saw (them) and said, <<Ake, our brother-in-law there went and speared our dinner.>>'
44. *Da iba sintuṅ' da ifa faṅa.*
and R-come put.down and R-go burn
'And (they) came and put it down and burned (it).'¹
45. *Ifaṅa sib da isaf.*
R-burn CMPLT and R-cut
'When (they) had burned (it), (they) cut (it).'
46. *Isaf, urubitan igi da iyus.*
R-cut egg-3pss there and R-abundant
'(They) cut (it open) and its eggs were plentiful.'

¹ It is burned to tighten up its skin for easier butchering (and to remove hair from animals that have it).

47. *Da ifab urubit igi ifa isar? garam igi santan ifa ntsup sib.*
 and R-distribute egg there R-go R-enough man there all R-go finish CMPLT
 ‘And (they) distributed the eggs until all the people there had received one.’ [Lit.: ‘until all the people there had run out first.’ i.e. They were out of people before they were out of eggs.]
48. *Da isaf fugai nidzum.*
 and R-cut crocodile true(meat)
 ‘Then (they) cut up the actual crocodile.’ [or ‘...the crocodile meat.’]
49. *Isaf sib ibinigi da itip ifab yatsur? urubit igi.*
 R-cut CMPLT R-like-that and R-repeat R-distribute R-follow egg there
 ‘When (they) had cut (it) up, (they) went around again distributing (it) following the eggs.’
50. *Ifab sib ibinigi da tuasaraja gan.*
 R-distribute CMPLT R-like-that and 3sg eat
 ‘After (they) had distributed (it) like that, they ate.’²
51. *Iga sib da tuasarajan impai.*
 R-eat CMPLT and 3sg R-sat
 ‘When (they) had eaten, they sat (there).’³
52. *Nan araja bawa? ibi arigi.*
 talk 3sg come-arrive R-like that.there
 ‘That’s the end of the story.’
53. *Da dzi ini bina?, Daŋki tsira? bijan. Nan iba ntsup.*
 and 1sg R-say like.this thanks big much talk R-come finish
 ‘So I say, Thank you very much. The story is finished.’

² Even though it is not stated explicitly, they would have cooked the meat before eating it.

³ Implied here is the idea that they sat and visited, telling stories, etc.

APPENDIX C

RELATIVE CLAUSES, COMPOUNDS, AND COMPLEMENT CLAUSES

More unambiguous examples of relative clauses:

- (C1) *Agu, [mama' finam [aru su-da gabun fa-da gamp marjan]_{RC}*
 2sg child female SPEC prepare-NMZ husband go-NMZ village another

igi]_{NP}, u watsuj' sanab rururuj-an...
 that 2sg NEG-follow road straight-NMZ

'You, young lady, who married into another village, you did not follow a good path/custom...' (Note: Marriage for a woman is expressed idiomatically as preparing food for a husband.)

- (C2) *[Garam [aru mpa-da yauj]_{RC} aga]_{NP} irim ujar aga.*
 man SPEC sit-NMZ shade that(far) R-build house that(far)

'That man right over there sitting in the shade built that house.'

- (C3) *[Iyam [aru naj-an da wap-a apu]_{RC} ugu]_{NP},*
 dog SPEC do-NMZ and steal-NMZ meat that(prev)

[garam [bugin-da iyam]_{RC} ugu]_{NP} yis funub.
 man hate-NMZ dog that.prev R-kill dead

'That dog who was always stealing meat, the man who hates dogs killed.'¹

- (C4) *Bi-da [garam [ataj'-a gamp]_{RC} igi]_{NP} itsaja² rungan garam utup iru',*
 like-NMZ men enter-NMZ village that R-see self-3pss men clan a.few

utup iru' bi-da bi' iru' da bits, da rib iru' da bits buj' aga nam.
 clan a.few like-NMZ blood two and one and 3pl two and one FUT IR-eat food

'If the men (who are) entering a/the village consider themselves to be men of a few clans, for example, three different families, then three men will eat food (and meat).'

¹ (C3) is in response to my attempt to elicit the passive sentence: 'The dog who is always stealing meat was killed by the man who hates dogs'.

² Normally the nominalized form of the verb *tsajan-da* would occur here in the conditional clause. However, because of the complexity of the sentence and perhaps to distinguish between the relative clause and the rest of the conditional clause, the speaker apparently felt it less awkward to use the realis form. Note that all other verbs in the conditional clause are nominalized.

- (C5) [Mama' [faja-da ujar]_{RC} aga]_{?³}]_{NP} irunt an aru-n.
 child burn-NMZ house that R-run 3pss way-3pss
 'The child who set the house on fire ran away.'
- (C6) [Dzi rainga' [bugin-dan iyam]_{RC} ugu]_{NP}
 1sg brother-1pss hate-NMZ dog that(prev)
 yis [iyam [wap-a tatar?]_{RC}]_{NP} funub.
 R-kill dog steal-NMZ chicken dead
 'My brother who hates dogs killed the dog who stole the chicken.'
- (C7) Warim bampij' da [sagat [tsinuf-a unas]_{RC}]_{NP}.
 IMP-give coconut to woman peel-NMZ sweet.potato
 'Give the coconut to the woman (who is) peeling sweet potato.'
- (C8) Wani [nan [idzuwai [u sau-da ruam in]_{RC}]_{NP}]_{NP}.⁴
 IMP-say talk what 2sg search.for-NMZ self-2pss 3obj
 'Tell (them) what [you need/lack].'

Examples of compound constructions:

- (C9) garam tip-a ujar⁵
 man repair-NMZ house
 'a carpenter'
- (C10) garam tip-a ujar funub
 man repair-NMZ house completely
 'a good carpenter' (or 'a man who is really good at building/fixing houses')

³ It is not always clear that the demonstrative modifies the head noun rather than a noun within the relative clause. The fact that the head noun in (C5) is definite suggests that the demonstrative belongs outside of the clause. The same is true in (C2), except that in that case the demonstrative could refer to both the man and the shade at the same time, since the man is sitting in the shade.

⁴ There is an extra level of embedding in (C8) because the inner NP describes what kind of talk. It is not 'what talk you lack' but 'talk of/about what you lack.'

⁵ S. Holzknrecht (1986:109) uses this same basic example in (C9), except hers is: *garam rim-a ujar*, from *rim-a* 'to put/make'. In (C10), the boundary between compounds and relative clauses starts to blur.

(C11) *garam yab-a bampij*⁶
 man climb-NMZ coconut

‘a man who climbs coconut palms’ (i.e. who likes to or is good at it)

More examples of complement clauses:

(C12) *I-tsanjan [intap dzufun²-a mpui, gum da mamai isi² isi² fain sib]*
 R-see earth bury-NMZ river garden and mountain small small some already

da uwayant i-wa² gin.
 and knowledge R-arrive 3obj

‘They saw that tons of earth covered rivers, gardens, and small mountains, and they realized what had happened.’ (my translation)

or, as the author translated it:

‘They saw ton after ton of earth covered small mountains, rivers, and gardens’

(C13) a. B-19Q *Wa-munti da dzi na-fa a-tsanja [fugai i-gij² mpui.]*
 IMP-stand and 1sg FUT-go IR-see crocodile R-lay water

‘Stand here and I will go see the crocodile laying in the water.’

b. B-20 ...*da i-tsanja fugai ru- gij²-an.* (ambiguous: RC or Comp. Cl.?)
 and R-see crocodile still lay-NMZ

‘and he saw the crocodile still laying there.’

Another ambiguous example of a relative or complement clause:

(C14) *da i-ri²an² i [simpup yai²-an mi²-a [gai wain*
 and R-hear Obj./Comp. drum cry.out-NMZ from-NMZ tree trunk

[araja mpa-da gin]_{RC} igi]_{NP} yab-an.]⁷
 3sg sit-NMZ 3obj Dem. go.up-NMZ

‘and he heard the sound of a kundu drum coming up from the trunk of the tree in which he sat.’ (my translation)

or

‘and he heard the kundu drum sound come up from the trunk of the tree where he sat on it.’ (author’s translation)

⁶ (C11) taken from S. Holzknecht (1986:109). This could be translated ‘a coconut climbing man’.

⁷ (C14) contains a clear example of a definite relative clause embedded within either another relative clause or a complement clause. Author’s translation suggests that it might be a complement clause.

REFERENCES

- Bartlett, Frederic C. 1932. *Remembering*. Cambridge: Cambridge University Press.
- Bickford, J. Albert. 1998. *Tools for Analyzing the World's Languages: Morphology and Syntax*. Dallas: SIL.
- Birner, Betty J. 1994. Information Status and Word Order: An Analysis of English Inversion. *Language* 70.233-59.
- Birner, Betty J. 1997. The Linguistic Realization of Inferrable Information. *Language and Communication* 17.133-147.
- Birner, Betty J., and Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins.
- Brown, Gillian, and George Yule. 1983. *Coherence in the Interpretation of Discourse*. *Discourse Analysis*, 223-56. Cambridge: Cambridge University Press.
- Burquest, Donald A. 2001. Interpreting Nominals in Hausa Narrative Discourse. *Text* 21.269-302.
- Chafe, Wallace L. 1974. Language and Consciousness. *Language* 50.111-33.
- Chafe, Wallace L. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. *Subject and Topic*, ed. by Charles N. Li, 27-55. New York: Academic Press.

- Chafe, Wallace L. 1987. Cognitive Constraints on Information Flow. *Coherence and Grounding in Discourse*, ed. by Russell S. Tomlin, 21-51. Amsterdam: John Benjamins.
- Chafe, Wallace L. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.
- Clark, Herbert H., and Susan E. Haviland. 1977. Comprehension and the Given-New Contract. *Discourse Production and Comprehension*, ed. by Roy O. Freedle, 1-40. Norwood, NJ: Apex Publishing Corporation.
- Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Cook, Walter A. 1998. *Case Grammar Applied*. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington.
- Cruse, D.A. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Currier, Constance Elizabeth. 1977. *Participant Reference in Bahasa Indonesia Narrative Discourse*. University of Texas at Arlington M.A. thesis. ms.
- Dempwolff, Otto. c.1928. *Analyse der Azera-Sprache*. ms. MS Book No. 786. Hamburg University Library.
- Dooley, Robert A., and Stephen H. Levinsohn. 2001. *Analysing Discourse: A Manual of Basic Concepts*. Dallas: SIL International.
- Donnellan, Keith. 1966. Reference and Definite Descriptions. *Philosophical Review* 75.281-304.
- Du Bois, John W. 1987. The Discourse Basis of Ergativity. *Language* 63.805-55.

- Fillmore, Charles J. 1997. *Lectures on Deixis*. (Reprint of 1971 Santa Cruz Lectures).
Stanford: CSLI Publications.
- Givón, T. (ed.) 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*.
Amsterdam: John Benjamins.
- Givón, T. 1991. Some Substantive Issues Concerning Verb Serialization: Grammatical vs.
Cognitive Packaging. *Serial Verbs: Grammatical, Comparative, and Cognitive
Approaches*, ed. by Claire Lefebvre, 137-84. Amsterdam: John Benjamins.
- Givón, T. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.
- Goris, Richard C., and Yukimi Okubo. (eds.) 1995. *Collins-Shubun English-Japanese
Dictionary*. Glasgow: Harper-Collins and Shubun International.
- Grice, H.P. 1975. *Logic and Conversation*. *Syntax and Semantics, Vol. 3: Speech Acts*, ed.
by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.
- Grimes, Barbara F. (ed.) 2000. *Ethnologue: Languages of the World, 14th Edition, CD-ROM
Version*. Dallas: SIL International.
- Grimes, Joseph E. 1975. *The Thread of Discourse*. The Hague: Mouton.
- Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. Cognitive Status and the
Form of Referring Expressions in Discourse. *Language* 69.274-307.
- Halliday, M.A.K. 1967. Notes on Transitivity and Theme in English, Part 2. *Journal of
Linguistics* 3.199-244.
- Halliday, M.A.K. 1970. *Language Structure and Language Function*. *New Horizons in
Linguistics*, ed. by John Lyons, 140-65. Harmondsworth, England: Penguin Books.

- Haviland, Susan E., and Herbert H. Clark. 1974. What's new? Acquiring New Information as a Process in Comprehension. *Journal of Verbal Learning and Verbal Behavior* 13.512-21.
- Holzkecht, H.A. 1974. Anthropological Research and Associated Findings in the Markham Valley of Papua New Guinea. *Research Bulletin* 15. Port Moresby: Department of Agriculture, Stock and Fisheries.
- Holzkecht, H.A. 1976. Amari-Adzera Social Structure. University of Queensland M.A. thesis. ms.
- Holzkecht, K.G. 1973a. The Phonemes of the Adzera Language. *Pacific Linguistics* A38.1-11.
- Holzkecht, K.G. 1973b. Morphophonemics of the Adzera Language. *Pacific Linguistics* A38.13-19.
- Holzkecht, K.G. 1973c. A Synopsis of Verb Forms in Adzera. *Pacific Linguistics* A38.21-8.
- Holzkecht, K.G. 1978. Dictionary: Adzera-English. ms.
- Holzkecht, Susanne. 1986. A Morphology and Grammar of Adzera (Amari Dialect), Morobe Province, Papua New Guinea. *Pacific Linguistics* A70.77-166.
- Holzkecht, Susanne. 1987. Why is smoke not a smoke? Language Taboo and Problems of Elicitation in the Languages of the Markham Family. *A World of Language: Papers Presented to Professor S.A. Wurm on his 65th Birthday*, ed. by D.C. Laycock and Werner Winter, 257-63. *Pacific Linguistics* C100. Canberra: Australian National University.

- Holzkecht, Susanne. 1988. Word Taboo and its Implications for Language Change in the Markham Family of Languages, Papua New Guinea. *Language and Linguistics in Melanesia* 18.43-69.
- Holzkecht, Susanne. 1989. *The Markham Languages of Papua New Guinea*. Pacific Linguistics C115. Canberra: Australian National University.
- Hopper, Paul, and Sandra A. Thompson. 1980. Transitivity in Grammar and Discourse. *Language* 56.251-99.
- Johnson, Morris. 1997. *Organized Phonology Data of Atsera*. Ukarumpa, PNG: SIL, ms.
- Jones, Larry B. 1983. Functions of First Mention References in Discourse. *Pragmatic Aspects of English Text Structure*, 49-75. Dallas: SIL.
- Lakoff, George. 1966. Stative Adjectives and Verbs in English. *Mathematical Linguistics and Automatic Translation*, ed. by Anthony G. Oettinger, I-1-16. Cambridge, MA: Harvard University Press.
- Lakoff, George. 1970. *Irregularity in Syntax*. New York: Holt, Rinehart, and Winston.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lefebvre, Claire. (ed.) 1991. *Serial Verbs: Grammatical, Comparative, and Cognitive Approaches*. Amsterdam: John Benjamins.
- Levinsohn, Stephen H. 1994. Field Procedures for the Analysis of Participant Reference in a Monologue Discourse. *Discourse Features of Ten Languages of West-Central Africa*,

- 109-21. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington.
- Longacre, Robert E. 1983. *The Grammar of Discourse*. New York: Plenum.
- Longacre, Robert E. 1989. *Joseph: A Story of Divine Providence: A Text Theoretical and Textlinguistic Analysis of Genesis 37 and 39-48*. Winona Lake: Eisenbrauns.
- Longacre, Robert E. 1995. Some Interlocking Concerns which Govern Participant Reference in Narrative. *Language Research* 31.697-714.
- National Statistical Office. 2002. *Papua New Guinea 2000 National Census: Community Profile System*. Port Moresby: N.S.O.
- Prince, Ellen F. 1981. Toward a Taxonomy of Given-New Information. *Radical Pragmatics*, ed. by P. Cole, 223-54. New York: Academic Press.
- Prince, Ellen F. 1992. The ZPG Letter: Subjects, Definiteness, and Information-status. *Discourse Description: Diverse Linguistic Analyses of a Fund-raising Text*, ed. by William C. Mann and Sandra A. Thompson, 295-325. Amsterdam: John Benjamins.
- Read, K.E. 1946. Social Organization in the Markham Valley, New Guinea. *Oceania* 17.93-118.
- Read, K.E. 1947. Effects of the Pacific War in the Markham Valley, New Guinea. *Oceania* 18.95-116.
- Read, K.E. 1949. The Political System of the Ngarawapum. *Oceania* 20.185-223.
- Rumelhart, David E. 1980. *Schemata: The Building Blocks of Cognition*. *Theoretical Issues in Reading Comprehension: Perspectives from Psychology, Linguistics, Artificial*

Intelligence, and Education, ed. by Rand J. Spiro, Bertram C. Bruce, and William F. Brewer, 33-58. Hillsdale, NJ: Lawrence Erlbaum.

Staley, William E. 1995. Referent Management in Olo: A Cognitive Perspective. University of Oregon PhD. dissertation. ms.

Velazquez-Castillo, Maura. 1996. The Grammar of Possession: Inalienability, Incorporation and Possessor Ascension in Guaraní. Amsterdam: John Benjamins.

Ward, Gregory. 1990. The Discourse Functions of VP Preposing. *Language* 66.742-63.

Ward, Gregory. 1999. A Comparison of Postposed Subjects in English and Italian. *Function and Structure: In Honor of Susumu Kuno*, ed. by Akio Kamio and Ken-ichi Takami, 3-21. Amsterdam: John Benjamins.

Ward, Gregory, and Betty J. Birner. 2001. Discourse and Information Structure. *The Handbook of Discourse Analysis*, ed. by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton, 119-37. Cambridge, MA: Blackwell.

BIOGRAPHICAL INFORMATION

David Edward Howard earned his M.A. in linguistics from The University of Texas at Arlington in 2002. He also has a certificate in biblical studies from Columbia Biblical Seminary, Columbia, South Carolina (1994), and a B.A. in classics from Bucknell University, Lewisburg, Pennsylvania (1993).

David and his wife, Susan, are members of the Summer Institute of Linguistics (SIL), and have worked in Papua New Guinea since 1998. David serves as a translation advisor and mentor for Adzera mother-tongue translators.